PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :
A01H 1/04, C12N 1/00, 5/00
C12N 15/00, C07H 21/04

A1

(11) International Publication Number: WO 89/12386

(43) International Publication Date: 28 December 1989 (28.12.89)

(21) International Application Number: PCT/US89/02729

(22) International Filing Date: 21 June 1989 (21.06.89)

(30) Priority data:

209,430	21 June 1988 (21.06.88)	US
262,442	25 October 1988 (25.10.88)	US
263,586	1 November 1988 (01.11.88)	US

(71) Applicant: CALGENE, INC. [US/US]; 1920 Fifth Street, Davis, CA 95616 (US).

(72) Inventors: WILKE-DOUGLAS, Melinda ; 4521 Barnett Ranch Road, Shingle Springs, CA 95686 (US). PERCHOROWICZ, John, T. ; 5631 N. Via Salerosa, Tucson, AZ 85715 (US). HOUCK, Catherine, M. ; 172 Carmel Court, Vacaville, CA 95686 (US). THOMAS, Bruce, R. ; 1606 Colusa, Davis, CA 95616 (US).

(74) Agent: RAE-VENTER, Barbara; Leydig, Voit & Mayer, 350 Cambridge Avenue, Suite 200, Palo Alto, CA 94306 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHODS AND COMPOSITIONS FOR ALTERING PHYSICAL CHARACTERISTICS OF FRUIT AND FRUIT PRODUCTS

(57) Abstract

Novel plants and methods for their preparation are provided which produce sucrase, particularly heterologous sucrase. The plants may have an increase in the soluble solids content of plant cells, particularly fruit cells as a result. DNA sequences encoding carbohydrate polymer synthesizing enzymes are identified, isolated and inserted into an expression cassette which may then be transformed into plant cells. The method finds particular use in conjunction with fruit-specific transcription for increasing production of carbohydrate polymers in fruit, particularly the apoplast.

METHODS AND COMPOSITIONS FOR ALTERING PHYSICAL CHARACTERISTICS OF FRUIT AND FRUIT PRODUCTS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. application Serial No. 265,586 filed November 1, 1988, which is a continuation-in-part of U.S. application Serial No. 262,442, filed October 25, 1988, which is a continuation-in-part of U.S. application Serial No. 209,430, filed June 21, 1988, which applications are hereby incorporated herein by reference.

INTRODUCTION

Technical Field

This invention relates to the use of carbohydrate polymers to alter the physical characteristics of fruit and fruit products.

Background

Of significant interest to the fruit-processing industry are the physical characteristics of fruit such as the soluble solids composition of ripe fruit. Soluble solids affect storage, handling, commercial processing and cooking of the fruit. Of particular interest is optimization of soluble solids characteristics of succulent fruit with regard to properties such as consistency and/or viscosity.

In various applications, fruit having increased soluble solids, increased or decreased viscosity, or increased or decreased consistency, is desired. Increased soluble solids may be a function of "sink strength," the ability of the plant to translocate carbon to the fruit. Increased fruit sink

activity is especially useful in fruit where high soluble solids content is desirable, such as tomatoes for processing into tomato-based products. Other important fruit characteristics include properties which relate generally to the cohesiveness of a homogeneous mixture of the fruit and the viscosity of the serum which remains after the pulp is removed through centrifugation.

Any modifications to fruit soluble solids, should not however have adverse effects on other desirable qualities or development of the plant or fruit. This includes consideration of the physiological consequences of the soluble solid, such as effects of osmotic potential, and/or any subsequent by-products on the plant, fruit and/or end-user, particularly, when the fruit is used for food. It is especially desirable that methods and compositions are developed to modify soluble solids which are safe for human consumption.

It is therefore of interest to obtain fruit having an altered carbohydrate content which may thus result in modified characteristics of the fruit or fruit products. In particular, methods and compositions for transforming a plant to produce fruit having modified solids characteristic are desirable. Furthermore, it is desirable not only to transform the plant cell, but also to obtain the altered carbohydrate composition at a particular stage in the growth of the plant, particularly fruiting. In addition, it is desirable that modification occur not only in a particular tissue, such as fruit, but also in a particular compartment of that tissue, such as the apoplast. It is therefore important to be able to express the carbohydrate polymer of interest using DNA sequences which provide for the desired regulation of expression in the plant host.

Relevant Literature

cDNA clones from tomato displaying differential expression during fruit development have been isolated and characterized (Mansson et al., Mol. Gen. Genet. (1985) 200:356-361; Slater et al., Plant Mol. Biol. (1985) 5:137-147). The studies have focused primarily on mRNAs which accumulate during fruit ripening. One of the proteins encoded by the ripening-specific cDNAs has been identified as polygalacturonase (Slater et al., Plant Mol. Biol. (1985) 5:137-147). A cDNA clone which encodes tomato polygalacturonase has been sequenced. Grierson et al., Nucleic Acids Research (1986) 14:8395-8603. The concentration of polygalacturonase mRNA increases 2000-fold between the immature-green and red-ripe stages of fruit development. This suggests that expression of the enzyme is regulated by the specific mRNA concentration which in turn is regulated by an increase in transcription. DellaPenna et al., Proc. Natl. Acad. Sci. USA (1986) 83:6420-6424. In vitro synthesis and processing of tomato fruit polygalacturanase has been described. DellaPenna et al. Plant Physiol (1988) 86:1057-1063. Mature plastid mRNA for psbA (one of the components of photosystem II) reaches its highest level late in fruit development, whereas after the onset of ripening, plastid mRNAs for other components of photosystem I and II decline to nondetectable levels in chromoplasts. Piechulla et al., Plant Mol. Biol. (1986) 7:367-376.

SUMMARY OF THE INVENTION

Novel compositions and methods are provided for modifying the soluble solids composition and physical characteristics of fruit and fruit products. The compositions include a novel DNA sequence encoding Leuconostoc mesenteroides. The methods involve transforming a plant cell of interest with an expression cassette functional in a plant cell comprising a trans-

criptional and translational initiation regulatory region, joined in reading frame 5' to a DNA sequence encoding an enzyme capable of polymerizing a cargohydrate comprising at least one hexose moiety to form a carbohydrate polymer and translational and transcriptional termination regions. To provide for transport of the enzyme to a specific cellular compartment, a DNA sequence encoding a transit peptide may be included in the expression cassette. Expression of the enzyme provides for an increase in the soluble solids composition of the fruit as a result of increased production of carbohydrate polymer. Of particular interest is the replacement of fruit soluble solids such as sucrose with carbohydrate polymers such as dextran or levan, depending upon the desired properties of the fruit or fruit products. Increased concentration of carbohydrate polymer in fruit can be used to alter not only the soluble solids composition of the fruit, but also characteristics, for example viscosity, of fruit products.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the integration of dextransucrase in tomato fruit metabolism.

Figure 2 shows restriction maps of dextransucrase (DS) clones pCGN1222, pCGN1223 and SALI partial digests of pCGN1222 (-16, -18, -1, and -69). pCGN1222-12 is a BamHI deletion of pCGN1222. Each plasmid was transformed into E. coli DH1 and extracts were assayed for DS activity (activity = cpm 14C-glucose incorporated into dextran): S = SalI; P = PstI; X = XbaI; E = EcoRI; B = BamHI; K = KpnI.

Figure 3 shows the restriction map of L. mesenteroides dextransucrase gene clone in the λ-fix vector (ds-18) and its derivatives in pUC19. Arrows indicate the direction of the lacZ promoter. The expression of dextransucrase activity of each clone in

E. coli is also shown. B, BamHI; E, EcoRI; H, HindIII; K, KpnI; P, PstI; R, EcoRV; S, SalI; Sc, SacI; Sp, SphI; X, XbaI.

Figure 4 shows the nucleotide sequence and flanking regions of the dextransucrase gene from L. mesenteroides. Numbering begins at the 5' end of the sequence. The deduced amino acids are those matching the four peptide sequence and the boxed amino acids are those matching the four peptide sequences deduced from amino acid sequence analysis. An arrow designates the proposed cleavage site for signal sequence removal. Putative promoter and ribosome binding site sequences are underlined. Inverted repeated sequences are underlined with broken arrows.

Figure 5 shows the amino acid analysis profile of the dextransucrase protein. Plotted is an average of three analyses. Cys and Trp percentages could not be determined under HCl cleavage conditions. Error bars indicate standard deviation. Protein sequence matches located within the translated DNA sequences are underlined.

Figure 6 shows a map of the 2A11 genomic clone. The transcriptional start site, the polyadenylation site, the start (ATG) and stop (TGA) sites and the position of the intron are indicated. The hatched region indicates the portion of the genomic clone that was used to make the tagged 2A11 constructions. The bottom portion shows the regions used to construct the 2A11 cassettes including the synthetic oligonucleotide used to insert restriction sites and reconstruct the 3' end.

Figure 7 shows the 2A11 cassette pCGN1241 containing the ampicillin resistance gene.

Figure 8 shows a strategy for construction of a PCR-generated dextransucrase construct.

Figure 9 shows a strategy of construction of 35S/dextransucrase and 2A11 dextransucrase binary

vectors.

Figure 10 shows a strategy for construction of a full length dextransucrase construct.

Figure 11 shows a strategy for cloning PG leader in front of a dextransucrase gene.

Figure 12 is a Northern analysis of leaves from plants transformed with B. subtilis levansucrase (LVS) gene.

(A) Two transformants (4-2, 8-2) contain transcripts that hybridize to the LVS probe, however only one transformant (8-2) is producing a full-length transcript (~2 kb).

(B) mas 5'-LVS-mas 3' construct in pCGN1205.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

In accordance with the subject invention, novel DNA sequences, DNA constructs, methods and compositions are provided which permit modification of the soluble solids composition of fruit by means of an increased synthesis of carbohydrate polymer from carbohydrates such as sucrose which comprise at least one hexose moiety. Plant cells are transformed with an expression cassette comprising a DNA sequence encoding a an enzyme such as a sucrase capable of polymerizing a carbohydrate such as sucrose to form a carbohydrate polymer. Timing of expression and/or tissue specificity, may be provided by the use of transcriptional regulatory regions having the desired expression specificity. The DNA sequence may also include a DNA sequence encoding a transit peptide recognized by the plant host to provide for targeting to a specific compartment within a tissue of interest, such as fruit apoplast.

Modification of the soluble solids composition of a fruit offers potential means to alter specific properties of the fruit. By "modification of the

soluble solids composition of fruit" is intended the replacement of sucrose, glucose and fructose or other soluble sugars with carbohydrate polymers. Of special interest are carbohydrate polymers which are not readily degraded. Although degradation of the polymer may be useful to achieve desired properties in some instances, most preferred is an investment in solids which will not be withdrawn later by degradation or other means. A related concern is the potential for an undesired accumulation of hexose monomers which may cause stress to the cell. Thus, polymers which are "unfamiliar" carbohydrate polymers, i.e., polymers which are not naturally found in the particular plant host of interest, may have special use because the plant is unlikely to possess an enzyme capable of degrading the unfamiliar polymer.

In addition to the effects of a modified carbohydrate profile from replacement of sucrose carbons with carbohydrate polymer carbons, the soluble solids of the fruit may be increased by the stimulation of added carbon flow to the fruit. If the draw of sucrose to carbohydrate polymer strengthens carbon-sink activity in the fruit, a corresponding increase in soluble solids may be achieved.

Considerations for use of a specific sucrase in fruit for the conversion of sucrose to a carbohydrate polymer include pH optimum of the enzyme, the availability of substrate and co-factors required by the enzyme. The sucrase enzyme of interest should have kinetic parameters compatible with the biochemical found in the host plant cell. For example, sucrase may have to compete for substrate with other enzymes, such as invertase. Analysis of the K_m and specific activity of the enzymes in question should be considered in determining the suitability of a sucrase for polymer production in a given host plant. Furthermore, in some cases, it may be desirable to target the enzyme to a

specific location. For example, in tomato it is desired to target the sucrase to the apoplast. The enzyme would thus need to be one which can function under conditions present in the apoplast.

For purposes of this invention, it is preferable to replace the native soluble solids with carbohydrate polymers which have, for example, different viscosity and/or consistency coefficients from the soluble solids of the fruit host. Replacement of tomato solids with levan, a viscous carbohydrate in pure solution, reduces gross viscosity and serum viscosity of tomato paste under simulated processing conditions. Replacement of tomato solids with dextran increases the gross viscosity of tomato paste. Thus the use of levansucrase and dextransucrase constructs, respectively, are of particular interest.

To provide for an increased expression of sucrase in a plant, a plant cell is transformed with an expression cassette which includes in the 5'-3' direction of transcription, a transcriptional and translational initiation region functional in a plant cell; a structural gene encoding a sucrase, preferably including a sequence encoding a transit peptide in reading frame at the 5'-terminus, where the transit peptide directs transfer of the sucrase to the apoplast; and a transcriptional and translational termination regulatory region. By sucrase is intended an enzyme capable of converting sucrose to a carbohydrate polymer. The initiation and termination regulatory regions are functional in the host plant cell and may be either homologous or heterologous to the host plant.

The sucrase gene may be derived from prokaryotic sources, for example, genes for dextransucrase, and other glycosyltransferases, which polymerize the glucose moiety of sucrose to form dextran, may be obtained from various species of Leuconostoc and

Streptococcus, including the soil bacterium L. mesenteroides. Genes for levansucrase, which catalyzes the conversion of sucrose to levan, a fructose polymer and free glucose, may be obtained from a variety of species of bacteria, for example, Aerobacter levanicum, Streptococcus salivarius and Bacillus subtilis. Preferentially the dextransucrase gene is derived from Leuconostoc mesenteroides and the levansucrase is derived from Bacillus subtilis. The DNA sequence encoding L. mesenteroides dextransucrase is provided. Dextransucrase derived from L. mesenteroides strain B-512F catalyzes the polymerization of branched dextran (95% α1->6 linkage and 5%-1->3 branched linkages) from sucrose. In contrast to Streptococcus strains, which typically produce several distinct glucosyltransferases, L. Mesenteroides B-512F produces only one glycosyltransferase, and at high levels. Comparison of the deduced amino acid sequence of dextransucrase with other sequenced proteins (data base) revealed partial homologies with two other glucoslytransferase proteins: gtfB from S. sobrinus (41.5%) and gtfB from S. mutants (41.7%).

Eukaryotic sources, including plant sources, such as Jerusalem artichoke and some grasses, may also find application. Examples of plant enzymes which also may be used include those involved in the synthesis of fructans from sucrose from a variety of sources (see, for example, Bhatia et al., Phytochem. (1979) 18:923-927; Edelman et al., New Phytol. (1968) 67:517-553; Henby et al., Phytochem. (1980) 19:1017-1020; Satyanarayana, Ind. J. Biochem. Biophys. (1976) 13:261-266; Shiomi, Agric. Biol. Chem. (1980) 44:603-614; Shiomi, Carbohydrate Research (1981) 96:281-292; Shiomi, Carbohydrate Research (1982) 99:157-169; and Singh, Phytochem. (1971) 10:2037-2039).

The structural gene for the sucrase may be obtained in a variety of ways. The gene may be synthesized in whole or in part, particularly where it is desirable to provide plant preferred codons. Thus, all or a portion of the open reading frame may be synthesized using codons preferred by the plant host. Plant preferred codons may be determined from the codons of highest frequency in the proteins expressed in the largest amount in the particular plant species of interest.

Methods for synthesizing sequences and bringing the sequences together are well established in the literature. Where a portion of the open reading frame is synthesized, and a portion is derived from natural sources, the synthesized portion may serve as a bridge between two naturally occurring portions, or may provide a 3'-terminus or a 5'-terminus. Particularly where the signal sequence and the open reading frame encoding the sucrase are derived from different genes, synthetic adapters commonly will be employed. In other instances, polylinkers may be employed, where the various fragments may be inserted at different restriction sites or substituted for a sequence in the polylinker.

The sucrase structural gene may be derived from cDNA, from chromosomal DNA or may be synthesized in whole or in part. For the most part, some or all of the sucrase structural gene will be from a natural source or genes substantially homologous to such sequences. In some situations it may be desirable to modify all or a portion of the codons, for example to enhance expression, by employing host-preferred codons. Methods for identifying sequences of interest have found extensive exemplification in the literature, although in individual situations, different degrees of difficulty may be encountered. Various techniques include the use of probes where genomic or cDNA libraries may be searched for complementary sequen-

ces. Where the structural gene to be inserted is derived from prokaryotic cells, it is desirable to minimize the 3' non-coding region of the prokaryotic gene. The substantial absence of this untranslated region can have a positive effect on the transcription, stability, and/or translation of the mRNA in the plant cells.

When the structural gene is other than a plant gene, in order to have expression of the gene in a plant cell transcriptional and translational initiation regulatory regions functional in a plant cell must be provided. Promoters and translation initiation signals functional in plant cells include those from genes which are present in the plant host or other plant species, for example the ribulose bisphosphate carboxylase small subunit transcriptional initiation region, for example from tobacco; those present in viruses such as the cauliflower mosaic virus (CaMV), for example the 35S transcriptional initiation region; and those associated with T-DNA such as the opine synthase transcriptional initiation regions, for example, octopine, mannopine, agropine, etc. Of particular interest is a transcriptional initiation region in a construct comprising two 355 promoters in tandem; fruit specific promoters, e.g. 2A11 (see co-pending application).

The regulatory regions may be homologous (derived from the original host) or heterologous (derived from a foreign source, or synthetic DNA sequence) to the plant host. In order to join the promoter(s) to the structural gene, the non-coding 5' region upstream from the structural gene may be removed by endonuclease restriction. Alternatively, where a convenient restriction site is present near the 5' terminus of the structural gene, the structural gene may be restricted and an adapter employed for linking the structural gene to a promoter region, where the adaptor provides for lost nucleotides of the structural

gene.

Transport of the sucrase into a particular cellular compartment may be accomplished by the use of a transit peptide to target a cellular compartment of interest, such as the apoplast. The transit peptide is inserted following the promoter sequence(s). The transit peptide and processing signal may be derived from any plant protein which is expressed in the cytoplasm and translocated to the cellular compartment of interest. The transit peptide can be identified by comparing the messenger RNA from the particular polypeptide with the mature product. The amino acid sequence absent from the mature protein and coded for by the messenger beginning at the initiation codon, usually a methionine, will normally be the transit sequence. Fragments from the natural transit sequence which retain their transport activity can also be used. The transit peptide is the sequence capable of directing a peptide joined to the transit peptide to the cellular compartment of interest and may be the whole wild-type transit peptide, a functional fragment thereof, or a functional mutant thereof. The full natural sequence is also included.

For the most part, the transit peptides from one plant are generally recognized by other plants. Thus the transit peptide may be native to or heterologous to the ultimate host in which the chimeric gene comprising the structural gene and transit peptide is introduced. Transit peptides may come from tomato or other fruits or vegetables. The transit peptide will usually have at least about 20 amino acids and not more than about 100 amino acids. Transit peptides of particular interest include leader sequences from genes encoding apoplast proteins produced in the cytoplasm, then translocated to the apoplast. Such proteins

include polygalacturonase (see Sheehy et al., Mol. Gen. Genet. (1987) 208:30-36), a cell wall enzyme implicated in the softening of fruit during ripening.

The termination region may be derived from the 3'-region of the gene from which the initiation region was obtained or from a different gene. The termination region may be derived from a plant gene, particularly the tobacco ribulose biphosphate carboxylase small subunit termination region; a gene associated with the Ti-plasmid such as the octopine synthase termination region; or the tml termination region.

In developing the expression cassette, the various fragments comprising the regulatory regions and open reading frame may be subjected to different processing conditions, such as ligation, restriction enzyme digestion, resection, in vitro mutagenesis, primer repair, use of linkers and adapters, and the like. Thus, nucleotide transitions, transversions, insertions, deletions, or the like, may be performed on the DNA which is employed in the regulatory regions and/or open reading frame. The expression cassette thus may be wholly or partially derived from natural sources, and either wholly or partially derived from sources homologous to the host cell, or heterologous to the host cell. Furthermore, the various DNA constructs (DNA sequences, vectors, plasmids, expression cassettes) of the invention are isolated and/or purified, or synthesized and thus are not "naturally occurring".

During the construction of the expression cassette, the various fragments of the DNA will usually be cloned in an appropriate cloning vector, which allows for amplification of the DNA, modification of the DNA or manipulation by joining or removing of sequences, linkers, or the like. Normally, the vectors will be capable of replication in at least a relatively high copy number in E. coli.

A number of vectors are readily available for cloning, including such vectors as pBR322, pUC series, M13 series, etc. The cloning vector will have one or more markers which provide for selection for transformants. The markers will normally provide for resistance to cytotoxic agents such as antibiotics, heavy metals, toxins, or the like. By appropriate restriction of the vector and cassette, and as appropriate, modification of the ends, by chewing back or filling in overhangs, to provide for blunt ends, by addition of linkers, by tailing, complementary ends can be provided for ligation and joining of the vector to the expression cassette or component thereof.

After each manipulation of the DNA in the development of the cassette, the plasmid will be cloned and isolated and, as required, the particular cassette component analyzed as to its sequence to ensure that the proper sequence has been obtained. Depending upon the nature of the manipulation, the desired sequence may be excised from the plasmid and introduced into a different vector or the plasmid may be restricted and the expression cassette component manipulated, as appropriate. The manner of transformation of E. coli with the various DNA constructs (plasmids and viruses) for cloning is not critical to this invention. Conjugation, transduction, transfection or transformation, for example, calcium chloride or phosphate mediated transformation, may be employed.

Depending upon the manner of introduction of the expression construct into the plant, other DNA sequences may be required. Commonly, the expression cassette will be joined to a replication system functional in prokaryotes, particularly E. coli, so as to allow for cloning of the expression cassette for isolation, sequencing, analysis, and the like. Included with the replication system will usually be one or more markers which may allow for selection in the host, the

markers usually involving biocide resistance, for example antibiotic resistance; heavy metal resistance; toxin resistance; complementation, providing phototrophy to an auxotrophic host; immunity; etc. Where the DNA will be microinjected into the host cell, a marker will usually be desirable which allows for selection of those cells in which the injected DNA has become integrated and functional. Thus, markers will be selected which can be detected in a plant host.

Alternatively, one may use agrobacteria as a vector for introducing DNA into a plant cell host. A first host vector is transformed with an extrachromosomal element or plasmid capable of maintenance in said first host, where said extrachromosomal element is characterized as being self-transmissable or capable of mobilization with a second helper extrachromosomal element; being capable of conjugal transfer to agrobacteria; containing predetermined DNA segments comprising foreign nucleic acid sequences, such as foreign DNA, encoding a peptide or other genetic information of interest capable of replication, transcription and usually translation in a host cell; and at least one DNA segment homologous with target DNA in agrobacteria, in said agrobacteria, either chromosomal or extrachromosomal; and optionally a marker for selection. Generally, the extrachromosonal elements comprise a vector having a replicon or replication system having a narrow host range. That is, the extrachromosomal elements are capable of maintenance and replication in a first host which is capable of conjugation with agrobacteria having the target DNA, but the vector is incapable of maintenance in the agrobacteria so that the agrobacteria is rapidly cured. In this way, in the absence of integration of at least a portion of the extrachromosomal composition, the extrachromosomal DNA will be lost in the agrobacteria. Thus, the DNA of the extrachromosomal element is rescued by integration with

a stable genetic element of the agrobacteria.

The replicon may come from a plasmid, virus, or other source which allows for stable maintenance in the first host. The first host will normally be a prokaryote, particularly E. coli. By employing a vector capable of replication in E. coli and other enteric hosts (narrow range), the vector will not be replicated and maintained in non-enteric Gram negative bacteria. The extrachromosomal element may therefore be a plasmid or virus (phage) or other DNA, which provides the requisite properties.

The extrachromosomal element may also have a mobilization locus, and be self-transmissable, or capable of conjugal transfer by trans-complementation with a helper plasmid having the necessary genes for transfer, e.g., tra. Conveniently, the extrachromosomal DNA and the helper plasmid will be initially employed in different cells, employing tripartite mating, or a variation thereof, rather than dual mating, so that in vitro introduction of both the extrachromosomal DNA and the helper plasmid into the first host cell prior to conjugal transfer is not required. The helper plasmid will have the necessary function for conjugal mobilizability (transfer), as well as a functioning replication system. Another characteristic is the presence of a DNA sequence in the extrachromosomal element of at least about 50bp which is at least substantially homologous with DNA present in the agrobacteria to provide for efficient integration.

The homologous DNA in the agrobacteria which serves as the recombination locus may be chromosomal or extra chromosomal. Of particular interest is the T-DNA region of the Ti plasmid, the region of the Ti plasmid between the left and right borders. The region of homology may include regions encoding proteins, regulatory regions, or non-encoding regions and need not have perfect homology. By having regions of at least

substantial homology with the target DNA on both sides of the sucrase expression cassette and in close proximity to the termini of the cassette DNA, the cassette may be inserted into the target host DNA with a minimum amount of the remaining DNA of the extrachromosomal element. Where it is desired to obtain normal plant cells and plants following transformation, the T-DNA sequences used to provide homology will not include the tumor genes, or the tumor genes will have been rendered inactive using means well known to those skilled in the art.

The region of homology will be of a sufficient size to provide for integration, and generally will be about 3000bp, more usually at least about 1000bp, conveniently 500bp, and may be 100bp or 50bp as to each of the fragments. The region of homology is generally present in two segments flanking the nucleic acid sequence of interest, the total number of bases of the segments generally not exceeding about 2000bp. The two segments generally not exceeding about 2000bp. The two segments may be the same or different, both as to size and sequence, or may be a wild type or mutant T-DNA sequence, interrupted by the intervening sequences; alternatively, one T-DNA sequence will suffice as described by Comai et al., Plasmid (1983) 10:21-30.

The use of T-DNA for transformation of plant cells has received extensive study ani is amply described in EPA Serial No. 120,516, Hoekema, In: The Binary Plant Vector System, Offsetdrukkerij Kanters B.V., Alblasserdam, 1985, Chapter V, Knauf, et al., "Genetic Analysis of Host Range Expression by Agrobacterium", In: Molecular Genetics of the Bacteria-Plant Interaction, Puhler, A. ed.,Springer-Verlag, NY, 1983, p. 245, and An et al., EMBO J. (1985) 4:277-284.

Conveniently, explants, cotyledons, or other plant tissue may be cultivated with A. tumefaciens or A. rhizogenes to allow for transfer of the expression

construct to the plant cells, the plant cells dispersed in an appropriate selective medium for selection, grown to callus, shoots grown and plantlets regenerated from the callus by growing in rooting medium. The agrobacterium host will contain a plasmid having the vir genes necessary for transfer of the T-DNA to the plant cells and may or may not have T-DNA. If the expression construct is to be inserted into the host cell by injection or electroporation, disarmed Ti-plasmids (lacking the tumor genes, particularly the T-DNA region) may be used.

Various techniques exist for determining whether the desired DNA sequences present in the plant cell are integrated into the genome and are being transcribed. Techniques such as the Northern blot can be employed for detecting messenger RNA which codes for the sucrase. In addition, the presence of expression can be detected in a variety of ways, such as assaying for enzyme activity or immunoassay for the protein product. A desired phenotype in this case is the presence of increased solids as polymers in the plant fruit, especially the presence of dextran or levan.

The cells which have been transformed may be grown into plants in accordance with conventional ways. See, for example, McCormick et al., Plant Cell Reports (1986) 5:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, identifying the resulting hybrid having the desired phenotypic characteristic. Two or more generations may be grown to ensure that the subject phenotypic characteristic is stably maintained and inherited and then seeds harvested for use to provide fruits with the new phenotypic property, particularly increased solids due to production of carbohydrate polymers, especially levan and dextran.

Modification of the carbohydrate profile of fruit may be used to alter the characteristics of fruit

products destined for human consumption. Thus, the safety of any food additive is an important consideration. Dextran is used clinically as a blood plasma extender. It formerly had U.S. "G.R.A.S." (generally recognized as safe) status as a food additive, but was dropped from the list solely due to disuse. One mole of glucose or fructose is released by the respective reactions of dextransucrase and levansucrase for each mole of sucrose consumed. Thus, even massive sucrose conversion should not "starve" the other metabolic needs of the fruit. Since the osmotic potential of one mole of sucrose is relatively equivalent to one mole of glucose or fructose monomer, the plant additionally will not be stressed and yield affected.

The following examples are offered by way of illustration and not by limitation.

EXPERIMENTAL

E. Coli 71-1.8; pCGN783 was deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Maryland, 20852, on December 23, 1988 and given Accession No. 67868.

Example 1

Effects of Carbohydrate Polymers on Viscosity of Tomato Paste

Effects on Viscosity

Polymers in water were added to tomato paste, which was then evaporated to achieve the desired level of total solids for each measurement. In all of these experiments, the polymer replaces part of the tomato solids. As a result, no change in the viscosity of the paste would be expected if the viscosity of the added polymer is equal to the average viscosity of the tomato solids that were replaced.

A. Levan

Modeling studies were done with levan from Aerobacter levanicum. This polymer is similar in structure to the levan produced by Streptococcus salivarius. Although levan is very viscous in pure solution, the polymer appears to break down under simulated tomato processing conditions. Thus, levan caused reductions in the gross viscosity and the serum viscosity of the tomato paste (Table 1).

Table 1

Effect Of Levan On Tomato Paste Viscosity

Hot break tomato juice	Bostwick (cm)	Ostwald (centisoles)	
control		4.5	597
+0.75% levan	5.8	198	

Bostwick consistency was measured at 10% total solids. Ostwald viscosity was measured on clear serum from paste at 6% total solids.

B. Dextran

Preliminary results from addition of dextran to tomato paste suggested that a dextran obtained from the Leuconostic mesertoroides B-512F increased Ostwald viscosity in both hot break paste and cold break paste (Table 2).

Table 2

Effect Of Dextran On
Tomato Paste Viscosity

Cold break tomato paste	Bostwick (cm)	Ostwald (centisoles)
control	7.5	152
+0.5% dextran	9.5	297
Hot break tomato paste	**Bostwick (cm)**	**Ostwald (centisoles)**
control	6.2	271
+0.5% dextran	6.5	430

Table 2 (cont'd)

Bostwick consistency was measured at 13% total solids (cold break paste) or at 10.8& total solids (hot break paste). Ostwald viscosity was measured on clear serum from paste at 6% solids.

C. Inulin

Inulin is a storage polymer produced by a number of higher plants which are grown for human consumption (e.g. Jerusalem artichoke and dandelion). No consistent effect on viscosity was observed when an inulin suspension was added to tomato paste at room temperature. The inulin used did not dissolve completely, however, unless heated to 80°C. Heating the inulin together with the tomato paste provided a more accurate model of how the inulin would be dispersed through the paste during processing. The temperature of the paste was then lowered for viscosity measurements. The result of this treatment was an increase in gross viscosity of the inulin-containing paste (Table 3).

Table 3

Effect Of Inulin On
Tomato Paste Viscosity

(a) Inulin suspension added to cold break tomato paste.

		Bostwick (cm)	Ostwald (centisoles)
Trial I	control	10.0	83
	+0.5% inulin	8.9	68
Trial II	control	11.0	71
	+0.5% inulin	12.9	65

(b) Inulin suspension added to hot break tomato paste.

		Bostwick (cm)	Ostwald (centisoles)
Trial I	control	6.0	270
	+0.5% inulin	7.0	271
Trial II	control	6.5	271
	+0.5% inulin	5.7	225

(c) Inulin heated together with cold break tomato paste.

		Bostwick (cm)	Ostwald centisoles
Trial I	control	11.0	71
Trial II	+0.5% inulin	6.5	79

Dextransucrase Enzymology

Dextransucrase activity was assayed under conditions chosen to approximate the environment of the tomato fruit apoplast as follows: 10 mM sucrose, 50 mM

glucose, 50 mM fructose, 3 mM $CaCl_2$, 50 mM sodium acetate pH 4.2, 0.02% NaN_3, 0.1 mg/ml Tween 80 and 25°C. (Daman et al. Plant Physiol. (1988) 87:731-736) Dextran polymer was isolated by ethanol precipitation, lyophilized and redissolved in 0.02% NaN_3. Dextran was assayed by the phenol-sulfuric acid method using glucose as the standard. Under these conditions, dextransucrase converted most of the sucrose into dextran polymers. The results were as shown in Table 4. Gel filtration and dextranase digestion demonstrated that the dextran synthesized in vitro had a structure indistinguishable from that of native dextrans. Only small amounts of the acceptor reaction products leucrose and isomaltose were produced. Thus, glucose and fructose were not readily used as acceptors and, at the concentrations found in tomato, they did not prevent the synthesis of high molecular weight dextran polymers.

Table 4

Dextransucrase: Reaction in vitro

	EtOH-precipitable carbohydrate formed within 72 hours	
reaction mixture	mg	% yield
complete (14 U/ml enzyme)	18	74
-minus enzyme control	none	0
-boiled enzyme control	none	0
-minus sucrose control	none	0

High molecular weight dextran polymer was also produced when dextransucrase and sucrose were added to serum from ripe tomato fruits (Table 5). This indicates that the tomato does not contain significant concentrations of any unknown substances that prevent

dextransucrase from producing high molecular weight dextran (e.g., inhibitors or acceptors).

Table 5

Dextran Synthesis in Tomato Fruit Serum[1]

	Reaction Mixture	High molecular weight carbohydrate (ug)
a	ripe fruit serum	0
b	serum+sucrose	74
c	serum+dextransucrase	40
d	serum+sucrose+dextransucrase	308
e	reaction mixture d + dextranase	80

[1]Sephacryl S-1000 gel filtration chromatography was performed on the reaction mixtures above. High molecular weight carbohydrate which eluted at the void volume had a molecular mass of >100,000 kDa. The 108 ug of high molecular weight carbohydrate in the original serum was substrated from the carbohydrate content of all reaction mixtures. High molecular weight carbohydrate was synthesized when sucrose and dextransucrase were added together (reaction mixture d). The newly synthesized dextran was hydrolyzed to low molecular weight products by treatment with dextranase enzyme (reaction mixture e).

In order to function in the tomato fruit apoplast, dextransucrase must compete for substrate with invertase (see Figure 1). The K_m for sucrose and the specific activity of these two enzymes were relatively close as shown in Table 6, below.

Table 6

Integration of Dextransucrase in Tomato Fruit Metabolism

	Km for sucrose (mM)	Specific activity (units/mg protein)
Tomato fruit invertase	6	320
Dextransucrase	14	170

Thus, dextransucrase can compete successfully for substrate and dextran polymers can be produced in the engineered fruits. At the other extreme, even massive dextran production should not "starve" the other metabolic needs of the fruit because one mole of fructose is released for every mole of sucrose consumed in dextran production.

Example 2
Isolation and Characterization of Dextransucrase Clones

Construction of L. mesenteroides DNA Library and Screening for Clones Expressing the Dextransucrase Gene

A Sau3A/Lambda-Fix (Stratagene) library was constructed, following the procedures supplied by the manufacturer, from L. mesenteroides DNA isolated from a late exponential phase culture of L. mesenteroides grown in GYC medium at 30°C. GYC medium has the following formulation: 10 g/l potassium phosphate (pH 6.7); 2 g/l yeast extract (Difco); 5 g/l casein hydrolysate (Gibco); 20 g/l glucose; and 5 ml/l salts solution (40 g $MgSO_4 \cdot H_2O$, 2 g NaCl, 2 g $FeSO_4 \cdot 7H_2O$, and 2 g $MnSO_4 \cdot H_2O$ per liter of water). DNA was partially digested with Sau3A to generate a random population of fragments averaging approximately 15-20 kb in size. Lambda-Fix is an EMBL vector modified so that the phage arms cannot ligate to themselves or to the stuffer fragment. In addition, the Sau3A partially-digested DNA is modified so that it cannot form multimers, eliminating the need to size fractionate the insert DNA. Sau3A partially-digested L. mesenteroides DNA was ligated into Lambda-Fix and packaged in vitro into viable phage particles. The Sau3A/Lambda-Fix library consists of approximately 100 genomes worth of L. mesenteroides DNA consisting of 10^5 plaque-forming units.

Recombinant phage ($5x10^3$) were screened for dextransucrase expression by the method of Gilpin et al. (Infect. Imm. (1985) 49:414) by plating with E. coli in Luria top agar (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl, 0.8% agar) over M-9 minimal salts agar (30 g/l 30g/l KH_2PO_4; 53g/l Na_2HPO_4; 5g/l NaCl; 10g/l NH_4Cl; 2g/l glucose; 1mM $MgSO_4$; 0.1mM $CaCl_2$), containing 0.4% sucrose; 0.1mg/ml dextran T-10; 1ug/ml thiamine; 8ug/ml leucine: 8µg/ml threonine, pH 7.5, and incubating at 30°C. The E. coli (strain C600) cannot utilize sucrose as a carbon source, therefore only weak growth of bacteria occurs in the top agar generating very small plaques. When the dextransucrase gene is present and expressed, dextran is produced and appears as a white halo around the plaque. Growth of E. coli is also stimulated around the plaque, presumably feeding off of the monosaccharides being released as dextrans are formed. Approximately 1% of the phage screened produced white halos of polymer around the plaques.

Dextran-producing plaques were isolated and partially purified. The clones were further analyzed by assaying for the incorporation of ^{14}C into polysaccharide polymer from sucrose labeled specifically in the glucose or fructose moiety. Polymer was picked from each of the putative dextransucrase clones growing in Luria top agar over sucrose minimal agar as described above and suspended in 100µl of DS buffer (200mM NaOAc, pH5.2, 0.1g/l Dextran T-9, 1.0g/l Tween 80, 1.0mM $CaCl_2$, 0.04% NaN_3). Half of the suspension (50µl) was then transferred to a reaction tube containing 0.1µCi of dried sucrose [glucose-^{14}C] with 50µl DS substrate (DS buffer, 6.8g/l sucrose) and the other half to a reaction tube containing sucrose [fructose-^{14}C] with 50µl DS substrate. The reactions were then incubated at 30°C for approximately 16 hrs. An aliquot of each

reaction (10-30µl) was then spotted onto Whatman No.1 filter paper. The filters were washed in three changes of methanol (10ml/filter/change) on a table top shaker for 10 minutes/change, then air dried and counted in 5ml scintillation fluid in a Beckman Scintillation Counter. Germaine et al, J. Dent. Res. (1974) 53: 1355-1360. All of the clones incorporated sucrose-derived ^{14}C-fructose and not ^{14}C-glucose into polymer indicating that the clones were expressing levansucrase rather than dextransucrase. In addition, reactivation of the enzyme produced after SDS-PAGE and PAS staining for carbohydrates (Miller and Robyt, Anal. Biochem. (1986) 156:357-363) revealed a protein with the mobility expected for levansucrase (116 kd).

Rescreening of the Lambda-Fix library was performed as described above except that the medium was first buffered to pH 6. The pH optimum of dextransucrase is pH 5.2. LB media is approximately pH 7.5 and increases to approximately pH 8 as the E. coli grow. At pH 8, dextransucrase retains only 10% of maximal activity, but this value is 80% at pH 6.0. Twenty clones producing halos were picked and assayed by the ^{14}C-labeling assay described above. One clone, DS-18, incorporated sucrose derived ^{14}C-glucose and was purified.

Example 3

Analysis of cDNA Clones

Determination of dextransucrase activity

DNA was isolated the DS-18 lambda clone and a partial restriction map was produced. A fragment containing the dextransucrase gene was subcloned as an XbaI fragment from DS-18 into pUC19 in both orientations (pCGN1222, pCGN1223). The plasmids were used to transform a DNA stabilizing strain of E. coli (DH1, recA). Maniatis et al., Molecular Cloning: A Laboratory Manual (Cold Springs Harbor, 1982).

Extracts were then prepared and assayed for DS activity by the sucrose [glucose-^{14}C] incorporation assay described above. Only one orientation of the insert (pCGN1222) resulted in expression of DS activity. Plasmids containing SalI partial deletions of pCGN1222 in DH1 were plated on SYC media (same as GYC media except sucrose is substituted for glucose) and grown at 30°C until colonies showed the large viscous colonies characteristics of dextran production at approximately nine days. The clones were further analyzed for DS activity by the sucrose [glucose-^{14}C] incorporation assay described above. Those colonies producing dextran also incorporated ^{14}C-glucose while colonies not producing dextran did not incorporate ^{14}C-glucose. Due to the orientation-dependent expression of DS described above, the 5' end of the DS gene was proposed to be at the left end of the pCGN1222 insert as shown in Figure 2.

Additional information was obtained by deletion of the BamHI fragment from pCGN1222 to form pCGN1222-12 (Fig. 2). The resulting plasmid retained the ability to incorporate ^{14}C-glucose from sucrose [glucose-^{14}C], although the activity appeared reduced. The remaining 5' 6.8kb fragment therefore presumably contains the majority of the DS gene, with perhaps some portion of the gene located beyond the BamHI site. Deletions were made to the pCGN1222 derivatives shown in Figure 2 to form the corresponding plasmids shown in Figure 3 using Promega's Erase-a-Base system. Progressive unidirectional deletions were made using exonuclease III which specifically digests DNA from 5' protruding or blunt ends while leaving 4 base 3' protruding ends intact. For example, plasmid pCGN1222-12 was cut with XbaI to generate a 5' protruding end and deletions were made approximately 3kb into the insert. The plasmid was also cut with SphI to generate a 3' protruding end thus protecting

the vector from digestion. Deletions were then transformed into E. coli (DH1), plated onto SYC media, and screened for DS activity by the sucrose [glucose-^{14}C] incorporation method described above. Plasmid DNA was isolated and restriction mapped from various colonies, both positive and negative for DS activity. The 5' boundary of the DS coding region was determined by comparing the restriction maps and DNA sequences of two clones, of pCGN2023, clones 5-1 and 5-7. Clone 5-1 does not incorporate ^{14}C-glucose. The 5' end of the insert maps 161bp 5' to the first PstI site and does not contain the first HindIII site. Clone 5-7 does incorporate ^{14}C-glucose and the 5' end of the insert maps 265bp 5' to the first PstI site and contains the first HindIII site (Fig. 3). Clone 5-1 does not contain the putative ATG initiation codon. Clone 5-7 contains an ATG (Met) codon which begins a 4.78kb open reading frame, and therefore is proposed to be the start site for DS translation. Clone 5-7 represents the shortest active fragment obtained.

DNA Sequencing

The DNA sequence of the dextransucrase (DS) gene was determined by the dideoxy-chain termination method of Sanger et al. (1977) using M13 single-stranded templates or double-stranded DNA (Chen and Seeburg) and the modified bacteriophage T7 polymerase (Sequences) (Tabor and Richardson). Fragments were subcloned into M13mp18 (Yannich-Perren et al., Gene (1985) 33:103-119) or a bluescribed M13+ chloramphenicol derivative (Stratagene, San Diego, CA) and nested deletions were created using the Erase-a-Base System (Promega Biotec). Synthetic oligonucleotide primers were synthesized on an Applied Biosystems model 380A DNA synthesizer to sequence through gaps and to determine the sequence of the complementary strand. Both strands of the Dex gene were completely sequenced. The sequence data were analyzed using the

IntelliGenetics Suite of molecular programs version S.3. The sequence is shown in Fig. 4.

Physical Characterization of the Gene Product of DS18

Purified DS18 plaques were plated (10^4 plaques/plate) onto M-9 sucrose (pH 6) in LB top agar (pH 6) and incubated at 30°C until polymer formation appeared (36-48 hrs). The top agar was scraped off and soaked in DS buffer (200mM NaOAc, pH5.2, 0.1g/l Dextran T-9, 1.0g/l Tween 80, 1.0mM $CaCl_2$, 0.04% NaN_3) to elute any proteins. The agar was then pelleted and the supernatant concentrated to approximately 50 µl. Sample buffer (0.05M Tris-HCl pH 6.8, 1.0% SDS, 5% β-mercaptoethanol, 10% glycerol, 0.005% bromophenol blue) was added and the samples were analyzed by SDS-PAGE. Silver staining revealed a protein band slightly larger than the size expected for dextransucrase (177 kd).

Comparison of DNA Sequence of DS18 Clone to Protein Sequence of Dextransucrase

Partially-purified dextransucrase protein (Miller et al., Carbohydrate Res. (1986) 147:119-133) was carboxymethylated and submitted to SDS-PAGE. A dextransucrase band was observed at 177 kd and a dextransucrase breakdown product formed a band at 158 kD. The breakdown product retains enzyme activity.

The majority of the protein preparation was concentrated (through Speed-Vac lyophilization) 10 fold and loaded onto four 0.5 mm 10% Laemmli SDS gels. Dextransucrase was separated from other contaminants by electrophoresis of gel for 60 min at 200v. The gel was stained for 2-4 min in 0.1% commassie blue, 47% methanol, 47% water, and 6% acetic acid; then destained in 50% methanol for 4-6 min. The clearly visible dextransucrase band in the upper portion of the gel was excised and soaked in 0.5% SDS, 1mM N-etnylmorpholine, pH 9.0 in preparation for electroelution.

Electroelution was carried out in an ISCO elution device. Gel chips were placed in the chamber assembled with 50kD molecular weight cutoff dialysis tubing and electroeluted for 120 min against 1 mM N-ethylmorpholine, pH 9.0 at 4°, 3 watts. Eluted sample was recovered from the concentrating well. The well was rinsed several times with additional buffer and pooled with the eluted sample to make a final volume of 8 ml. Subsequent SDS-PAGE analysis established the presence of a single molecular weight species at approximately 177kD.

Aliquots (5% each) were taken and dried in preparation for amino acid hydrolysis. One sample was performic acid oxidized to enable identification of cysteine and methionine residues. Samples were then subjected to vapor phase HCl/phenol hydrolysis for 24 hr, 110°C in vacuo. Amino acid composition was determined by sample application to a Beckman 6300 amino acid analyzer. (See Figure 5.)

Peptides were generated from one-third of the remaining dextransucrase to facilitate determination of protein sequence. Chemical cleavage with cyanogen bromide was chosen to optimize the number of peptides generated. The protein was lyophilized to dryness, resuspended in 98% formic acid, then adjusted to 70% formic acid by adding water. β-mercaptoethanol was added to 1%. Cyanogen bromide crystals were added to a concentration of approximately 20-30 mg/ml. The reaction proceeded for 24 hr at room temperature, in the dark at which time a sample was analyzed by SDS gel electrophoresis. The remaining digest was size fractionated over a HPLC TSK 2000sw sizing column in 50% acetonitrile, 0.1% TFA.

Two fractions from this sizing column were each partially dried down and brought up to 100 µl in 0.1% TFA. Further fractionation of the peptides contained in these size cuts was achieved through HPLC

chromatography on a narrow bore, reverse phase C8 column using an acetonitrile/H_2O gradient. Material recovered from this column served as the sample for protein sequencing.

Based on a comparison of the readable sequence to the DNA sequence, four peptides were identified in the sequencing sample. The four peptide sequences can be matched to the DNA sequence. The four peptides read as follows:

1) NFGTITAN-A-A-F
2) VDRSNDSTENEAIPNYSFV-R-D-V
3) D-Y
4) -E/GYAFLEFLQ/VAL

(dashes indicate the amino acid at the site was not identified, slashes indicate an either/or situation). The location of the peptide sequences in the translation of the DNA sequence is shown in Figure 5.

Example 4
Construction of Levansucrase Expression Cassette

Plasmid pLS8 containing the *B. subtilis* levansucrase gene was obtained from P. Gay (Gay *et al.*, *J. Bacteriol.* (1983) 153:1424-1431). The *Rsa*I-*Eco*RI fragment containing the 5'-end of the gene and the *Eco*RI-*Eco*RV fragment containing the 3'-end of the gene were isolated from polyacrylamide and agarose gels, respectively, and ligated into the *Sma*I site of Bluescribe M13 (minus). The resulting clones, pCGN498 and pCGN499 have the levansucrase gene inserted into the Bluescribe polylinker in opposite directions.

Insertion of Levansucrase Gene into the mas Cassette

pCGN1047 is an expression cassette consisting of the 5' and 3' non-coding regions of the mannopine

synthase (mas) gene (position 20804 to 20126) of Barker et al., Plant Mol. Biol. (1983) 2:335-350) with a poly-linker between the 5' and 3' regions. The insert from pCGN499 was excised as a Bam-Sst fragment and ligated into Bam-Sst cut pCGN1047. The resulting plasmid, pCGN1201, contained the levansucrase coding region between the 5' and 3' regions of the mas promoter (Figure 12).

Transfer of mas-Levansucrase Construction into Broad Host Range Binary Plasmid

The binary plasmid used for transfer of the mas-levansucrase construction to Agrobacterium was pCGN783 described in co-pending U.S. application serial number 188,361 filed April 29, 1988 which is hereby incorporated by reference.

(a) Construction of pCGN783

pCGN783 is a binary plasmid containing the left and right T-DNA borders of A. tumefaciens octopine Ti-plasmid pTiA6 (Currier and Nester, J. Bacteriol. (1976) 126:157-165) the gentamicin resistance gene of pPH1J1 (Hirsch et al., Plasmid (1984) 12:139-141), the 35S promoter of cauliflower mosaic virus (CaMV) (Gardner et al., Nucleic Acid Res. (1981) 9:1871-1880); the kanamycin resistance gene of Tn5 (Jorgensen, Mol. Gen. (1979) 177:65); and the 3' region from transcript 7 of pTiA6 (Currier and Nester, supra (1976)).

(b) Construction of pCGN587

The HindIII-SmaI fragment of Tn5 containing the entire structural gene for APH3'II (Jorgensen et al., Mol. Gen. (1979) 177:65), was cloned into pUC8 (Vieira and Messing, Gene (1982) 19:259), converting the fragment into a HindIII-EcoRI fragment, since there is an EcoRI site immediately adjacent to the SmaI site. The PstI-EcoRI fragment containing the 3' portion of

the APH3'II gene was then combined with an EcoRI-BamHI-SalI-PstI linker into the EcoRI site of pUC7 (pCGN546W). Since this construct does not confer kanamycin resistance, kanamycin resistance was obtained by inserting the BglI-PstI fragment of the APH3'II gene into the BamHI-PstI site (pCGN546X). This procedure reassembles the APH3'II gene, so that EcoRI sites flank the gene. An ATG codon was upstream from and out of reading frame with the ATG initiation codon of APH3'II. The undesired ATG was avoided by inserting a Sau3A-PstI fragment from the 5' end of APH3'II, which fragment lacks the superfluous ATG, into the BamHI-PstI site of pCGN546W to provide plasmid pCGN550. The EcoRI fragment of pCGN550 containing the APH3'II gene was then cloned into the EcoRI site of pUC8-pUC13 (K. Buckley supra (1985)) to give pCGN551.

Each of the EcoRI fragments containing the APH3'II gene was then cloned into the unique EcoRI site of pCGN451, which contains an octopine synthase cassette for expression to provide pCGN548 (2ATG)) and pCGN552 (1ATG). The plasmid pCGN451 having the ocs 5' and the ocs 3' in the proper orientation was digested with EcoRI and the EcoRI fragment from pCGN551 containing the intact kanamycin resistance gene inserted with EcoRI site to provide pCGN552 having the kanamycin resistance gene in the proper orientation. This ocs/KAN gene was used to provide a selectable marker for the trans type binary vector pCGN587.

The 5' portion of the engineered octopine synthase promoter cassette consists of pTiA6 DNA from the XhoI at bp 15208-13644 (Barker et al., supra (1983)), which also contains the T-DNA boundary sequence (border) implicated in T-DNA transfer. In the plasmid pCGN587, the ocs/KAN gene from pCGN552 provides a selectable marker as well as the right border. The left boundary region was first cloned in M13mp9 as a HindIII-SmaI piece (pCGN502) (base pairs 602-2212) and recloned as a

KpnI-EcoRI fragment in pCGN565 to provide pCGN580. As described above, pCGN565 is a cloning vector based on pUC8-Cm, but containing pUC18 linkers. pCGN580 was linearized with BamHI and used to replace the smaller BglI fragment of pVCK102 (Knauf and Nester, Plasmid (1982) 8:45), creating pCGN585. By replacing the smaller SalI fragment of pCGN585 with the XhoI fragment from pCGN552 containing the ocs/KAN gene, pCGN587 was obtained.

(c) Construction of pCGN739 (Binary Vector)

To obtain the gentamicin resistance marker, the resistance gene was isolated from a 3.1 kb EcoRI-PstI fragment of pPHIJI (Hirsch et al., Plasmid (1984) 12:139-141) and cloned into pUC9 (Vieira et al., Gene (1982) 19:259-268) yielding pCGN549.

The pCGN549 HindIII-BamHI fragment containing the gentamicin resistance gene replaced the HindIII-BglII fragment of pCGN587 creating pCGN594.

The pCGN594 HindIII-BamHI region which contains an ocs-kanamycin-ocs fragment was replaced with the HindIII-BamHI polylinker region from pUC18 (Yanisch-Perron, Gene (1985) 33:103-119) to make pCGN739.

(d) Construction of 726c (1 ATG-Kanamycin-3' region)

pCGN566 contains the EcoRI-HindIII linker of pUC18 (Yanisch-Perron, ibid) inserted into the EcoRI-HindIII sites of pUC13-Cm (K. Buckley, Ph.D. Thesis, University of California, San Diego, 1985). The HindIII-BglII fragment of pNW31c-8, 29-1 (Thomashow et al., Cell (1980) 19:729) containing ORF1 and 2 (Barker et al., Plant Mol. Biol. (1984) 2:335-350) was subcloned into the HindIII-BamHI sites of pCGN566 producing pCGN703.

The Sau3A fragment of pCGN703 containing the 3' region of transcript 7 from pTiA6 (corresponding to bases 2396-2920 of pTi15955 (Barker et al., supra (1984)) was subcloned into the BamHI site of pUC18 (Yanisch-

Perron et al., supra (1985)) producing pCGN709.

The EcoRI-SmaI polylinker region of pCGN709 was replaced with the EcoRI-SmaI fragment from pCGN587 (see 6.1(a), supra) which contains the kanamycin resistance gene (APH3'II) producing pCGN726.

The EcoRI-SalI fragment of pCGN726 plus the BglII-SalI sites of pUC8-pUC13-cm (chloramphenicol resistant, K. Buckley, Ph.D. Thesis, University of California, San Diego, 1985) producing pCGN738. To construct pCGN734, the HindIII-SphI site of M13mp19 (Norrander et al., Gene (1983) 26:101-106). Using an oligonucleotide corresponding to bases 3287 to 3300, DNA synthesis was primed from this template. Following S1 nuclease treatment and HindIII digestion, the resulting fragment was cloned into the HindIII-SmaI site of pUC19 (Yanisch-Perron et al., supra (1985)). The resulting EcoRI to HindIII fragment of pTiA6 (corresponding to bases 3390-4494) was cloned into the EcoRI site of pUC8 (Vieira and Messing, supra (1982)) resulting in pCGN734. pCGN726c is derived from pCGN738 by deleting the 900 bp EcoRI-EcoRI fragment.

(e) Construction of pCGN167

pCGN167 is a construct containing a full length CaMV promoter, 1 ATG-kanamycin gene, 3' end and the bacterial Tn903-type kanamycin gene. MI is an EcoRI fragment from pCGN550 (see construction of pCGN587) and was cloned into the EcoRI cloning site in the 1 ATG-kanamycin gene proximal to the polylinker region of M13mp9. See copending Application Serial No. 920,574, filed October 17, 1986, which disclosure is incorporated herein by reference.

To construct pCGN167, the AluI fragment of CaMV (bp 7144-7735) (Gardner et al., Nucl. Acids Res. (1981) 9:2871-2888) was obtained by digestion with AluI and cloned into the HincII site of M13mp7 (Vieira, Gene (1982) 19:259) to create C614. An EcoRI digest of C614

produced the EcoRI fragment from C614 containing the 35S promoter which was cloned into the EcoRI site of pUC8 (Vieira et al., Gene (1982) 19:259) to produce pCGN146. To trim the promoter region, the BglII site (bp 7670) was treated with BglII and Bal31 and subsequently a BglII linker was attached to the Bal31 treated DNA to produce pCGN147.

pCGN148a containing the promoter region, selectable marker (KAN with 2 ATGs) and 3' region was prepared by digesting pCGN528 (see below) with BglII and inserting the BamHI-BglII promoter fragment from pCGN147. This fragment was cloned into the BglII site of pCGN528 so that the BglII site was proximal to the kanamycin gene of pCGN528.

The shuttle vector used for this construct, pCGN528, was made as follows. pCGN525 was made by digesting a plasmid containing Tn5 which harbors a kanamycin gene (Jorgenson et al., Mol. Gen. (1979) 177:65) with HindIII-BamHI and inserting the HindIII-BamHI fragment containing the kanamycin gene into the HindIII-BamHI sites in the tetracycline gene of pACYC184 (Chang and Cohen, J. Bacteriol. (1978) 134: 1141-1156). pCGN526 was made by inserting the BamHI fragment 19 of pTiA6 (Thomashow et al., Cell (1980) 19:729-739) into the BamHI site of pCGN525. pCGN528 was obtained by deleting the small XhoI fragment from pCGN526 by digesting with XhoI and religating.

pCGN149a was made by cloning the BamHI kanamycin gene fragment from pMB9KanXXI into the BamHI site of pCGN148a. pMB9KanXXI is a pUC4K variant (Vieira and Messing, Gene (1982) 19:259-268) which has the XhoI site missing but contains a functional kanamycin gene from Tn903 to allow for efficient selection in Agrobacterium.

pCGN149a was digested with BglII and SphI. This small BglII-SphI fragment of pCGN149a was replaced with the BamHI-SphI fragment from MI isolated by digestion with BamHI and SphI. This produces pCGN167.

(f) Construction of pCGN766c (35S promoter-3' region)

The HindIII-BamHI fragment of pCGN167 containing the CaMV-35S promoter, 1 ATG-kanamycin gene and the BamHI fragment 19 of pTiA6 was cloned into the BamHI-HindIII sites of pUC19 (Norrander et al., supra (1985); Yanisch-Perron et al., supra (1985)) creating pCGN976.

The 35S promoter and 3' region from transcript 7 was developed by inserting a 0.7 kb HindIII-EcoRI fragment of pCGN976 (35S promoter) and the 0.5 kb EcoRI-SalI fragment of pCGN709 (transcript 7:3' for construction see supra) into the HindIII-SalI sites of pCGN566 creating pCGN766c.

(g) Final Construction of pCGN783

The 0.7 kb HindIII-EcoRI fragment of pCGN766c (CaMV-35S promoter) was ligated to the 1.5 kb EcoRI-SalI fragment of pCGN726c (1-ATG-KAN-3' region) into the HindIII-SalI sites of pUC119 (J. Vieira, Rutgers University, New Jersey) to produce pCGN778. The 2.2 kb region of pCGN778, HindIII-SalI fragment containing the CaMV 35S promoter (1-ATG-KAN-3' region) replaced the HindIII-SalI polylinker region of pCGN739 to produce pCGN783.

The mas-levansucrase cassette pCGN1201 was inserted into the unique SalI site of pCGN783 to create plasmid pCGN1205. This plasmid in E. coli C2110 is conjugated into Agrobacterium strain PC2760 (Ooms et al., Plasmid (1982) 7:15-29; Hoekema et al., Nature (1983) 303:179-181; European Patent Application 84-200239; 2424283) which contains a disarmed Ti-plasmid, using gentamicin selection.

Example 5
Construction of Dextransucrase Expression Cassettes

Expression of Cassettes Generally

The dextransucrase gene from pCGN1222 was cloned through a series of steps into pCGN2306 (Figure 8). The dextransucrase gene was then removed from pCGN2306 as an EcoRV fragment and cloned into the respective promoter cassettes (pCGN2305, having the 35S CaMV promoter and pCGN1241, having the 2A11 promoter). The resulting constructs were then cloned into a binary cassette (pCGN1547) in both orientations, yielding a total of 4 constructs, pCGN2318 and 2319 and pCGN2310 and 2311 (Figure 9).

2A11 Promoter Cassette

The design of the 2A11 cassette (pCGN1241) is shown in Figure 7. The cassette contains 3.8 kb of DNA 5' of the transcriptional start site and the entire 3' region (from the TGA stop codon to a site 2.0 kb 3' of the poly A addition site) of the 2A11 gene. A map of the 2A11 gene is shown in Figure 6, showing the restriction sites and indicating (below the representation of the gene) the regions of the 2A11 gene used to construct the 2A11 cassette. The 2A11 cassette was constructed as follows.

The 5' end of the 2A11 cassette was constructed starting with an EcoRI subclone genomic clone as described in application PCT/US88/01811 cloned into the EcoRI site of Bluescript (+) (Stratagene) resulting in pCGN1288. This clone contains sequences from the EcoRI site at position 1651 in the intron of the 2A11 gene to the EcoRI site located 2.5 Kb upstream of the XhoI site at position 1 of the sequenced region (see Figure 6). The XhoI fragment from position 1 of the sequenced region to the XhoI site in the Bluescript

polylinker was deleted creating plasmid pCGN2004 which contains the 2A11 region from position 1 to position 1651. The coding region of 2A11 was deleted by treating this plasmid with ExonucleaseIII/S1 (using the commercially available Erase-a-Base Kit (Promega Biotec)) and sequencing deletion plasmids until one was found which had the coding region deleted to position 1366. The resulting plasmid, pCGN1251, had the genomic region from the XhoI site (position 1) to position 1366. The EcoRI fragment of pCGN1288 was then transferred to a chloramphenicol resistant plasmid vector, pCGN2015, to make pCGN1231. pCGN2015 is a Cm resistant derivative of the Bluescript plasmid. A BstEII/BamHI fragment of pCGN1251 was then transferred into BstEII/BamHI digested pCGN1231 to make pCGN1235 which contains the region from the EcoRI site (2.5 kb upstream of the sequenced region) to position 1366 of the sequenced region flanked by the Bluescript polylinker in a Cm resistant vector.

The 3' end of the 2A11 cassette was constructed from pCGN1273 (described in application PCT/US88/01811) by digesting the plasmid with PvuI and EcoRI, isolating the 2249 bp insert (from position 2402 to 4653), ligating with a double-stranded oligonucleotide containing the sequence shown in Figure 6 from the BamHI sticky end to a PvuI sticky end into a Bluescript vector which had been digested with BamHI and EcoRI. The resulting plasmid, pCGN1238 contains the 3' end of the 2A11 gene from the stop codon at position 2381 to the EcoRI site at position 4653.

A cassette containing the 5' and 3' regions of the 2A11 gene was constructed by ligating the BamHI to EcoRV insert of pCGN1238 into pCGN1235 which had been digested with BamHI and XbaI (the XbaI site having been filled in with Klenow polymerase to make a blunt-ended fragment). The resulting plasmid, pCGN1240, has the 5' end of the 2A11 gene from the EcoRI site 2.5 kb

upstream of the XhoI site (position 1) to position 1366 (which is located between the transcriptional initiation site of the 2A11 gene and the ATG), followed by a polylinker region (sequence given in Figure 6) with sites for SmaI, BamHI, PstI and SalI which can be conveniently used to insert genes followed by the 3' region from position 2381 to 4653. The plasmid backbone of pCGN1240 is the Bluescript Cm plasmid described above. A more convenient version has the EcoRI of 1240 excised and inserted into a Bluescript vector with an altered polylinker from the SacI site to the KpnI site with a synthetic polylinker with the following sequence:

AGCTCGGTACCGAATTCGAGCTCGGTAC to create a polylinker with the following sites: SacI-KpnI-SacI-KpnI. The EcoRI insert of pCGN1240 was inserted into pCGN1240 to make pCGN1241 (Figure 7).

35S Promoter Cassette

The 35S CaMV cassette (pCGN2305) contains DNA 5' of the transcriptional start site of the 35S CaMV gene of the 3' region of the Tr7 gene. It was constructed as follows. pCGN1410 (Sheehy et al., (1988) Proc. Natl. Acad. Sci. (USA), 85:8805-8809) is digested with EcoRI, and the ends filled in with DNA polymerase I. A phoshorylated XbaI linker from Pharmacia, 5'-pd[TCTAGA]-3' is ligated in. The resulting plasmid was called pCGN2304. pCGN2304 was then digested with HindIII, the ends filled in with DNA polymerase I, and a phosphorylated XbaI linker ligated in. The resulting plasmid was called pCGN2305.

Preparation of Necessary Constructs

pCGN2302. The 3' end of the gene was subcloned from pCGN1222 (see Example 3) as a 3.1 Kb SalI/EcoRV fragment into pCGN2015 to give pCGN2302. The plasmid pCGN2015 was constructed by digesting

pCGN565 with HhaI blunting with mung bean nuclease and ligating with EcoRV digested Bluescript KSM13 (Stratagene, San Diego, CA) to create pCGN2008. pCGN565 is a cloning vector based on pUC8-pUC13-cm (K. Buckley (1985), supra) but containing pUC18 linkers (Yanisch-Perron et al., Gene (1985) 53:103-119).

pCGN2008 was digested with EcoRI and HindIII, blunted with Klenow. The 1156 bp chloramphenicol fragment was isolated. Bluescript KSM13+ (Stratagene) was digested with DraI and the 2273 bp fragment is isolated and ligated with the pCGN2008 chloramphenicol fragment creating pCGN2015.

pCGN2306. The remaining 1.6 kb 5' end of the dextransucrase gene was generated using the polymerase chain reaction method, Tag polymerase (Gene Amp kit) and the DNA thermal cycler (Perkin-Elmer Cetus) using pCGN2023 as a template. pCGN was made by a BamHI deletion of pCGN1222. Two oligonucleotides, Dex02 (42-mer with the added restriction sites, XhoI and EcoRV upstream of the ATG codon, plus sequence complementary to the ATG region) and Dex03 (a 26-mer complimentary to the SalI region) were used as primers.Dex02: 5'-GGTTATCTCGAGGATATCATGCCATTTACAGAAAAGTAATG-3'
Dex03: 5'-CAAATCAGCATCCACATTGTCGACTG-3'

The resulting 1.6 kb fragment contains the 5' end of the gene, from the ATG start codon through the SalI site with an additional SkoI site and an EcoRV site upstream of the ATG. The 1.6 kb fragment was then cloned in pCGN2303 as an XhoI/SalI fragment, in front of the 3' end of the gene giving plasmid pCGN2306.

pCGN1547. pCGN1547 (described in co-pending U.S. patent application Serial No. 329,018, (which disclosure is hereby incorporated by reference) is a binary plant transformation vector containing the left and right T-DNA borders of Agrobacterium tumefaciens octopine Ti-plasmid pTiA6 (Currier and Nester, J. Bact. (1976) 126:157-165), the gentamycin resistance gene of

pPH1JI (Hirsch and Beringer, Plasmid (1984) 12:139-141), an Agrobacterium rhizogenes Ri plasmid origin of replication from pLJbB11 (Jouanin et al., Mol. Gen. Genet. (1985) 201:370-141), the mas promoter region and mas 3' region of pTiA6 with the kanamycin resistance gene of Tn5 (Jorgensen et al., Mol. Gen. Genet. (1979) 177:65), a ColE1 origin of replication from pBR322 (Bolivar et al., Gene 2:95-113), and a lacZ' screenable marker gene from pUC18 (Yannisch-Perron et al., Gene 33:103-119).

Final Construction Steps

The entire dextransucrase gene was removed from pCGN2306 as an EcoRV fragment and cloned as a blunt ligation into the SmaI sites of the promoter cassettes pCGN2305, giving plasmid pCGN2309, and pCGN1241, giving pCGN2307. Each cassette was cloned into the binary vector pCGN1547 as an XbaI fragment for the 35S cassette to create pCGN2318 and pCGN2319, and as a KpnI fragment fro the 2A11 cassette to create8 pCGN2310 and 2311. This allows for constructs in both orientations for each cassette. The completed binaries were then submitted for cocultivation. Agrobacterium tumefaciens strain 2760 (also known as LBA 4404, Hoekema et al., Nature (1983) 303:179-180) were transformed with the receptive binary using the method of Holsters et al. Mol. Gen. Genet. (1978) 163:181-187. The transformed binary was then used in the cocultivation of plants.

Example 6

Preparation of Transgenic Plants

Feeder plates were prepared by pipetting 0.5 ml of an eight day old suspension of Nicotiana tabacum cv xanthi cell suspension culture (~10^6 cells/ml) onto 0.8% agar medium, containing MS salts, myo-inositol (100 mg/l), thiamine-HCl (1.3 mg/l), sucrose (30 g/l), potassium acid phosphate (200 mg/l) 2,4-D (0.2 mg/l),

and kinetin (0.1 mg/l) (pH 5.5). The feeder cells were prepared at least 24 hours prior to use. A #1 Whatman sterile filter paper (Whatman Ltd, Maidstone, England) was placed on top of the tobacco feeder cells after the cells had been growing for at least 24 hours.

Agrobacteria containing the plasmid of interest were grown on AB medium (AB salts K_2HPO_4 3 gm/l, NaH_2PO4 H_2O 1.15 g/l, NH_4Cl 1 g/l, glucose 5 g/l, $FeSO_4$ 0.25 mg/l, $MgSO_4$ 0.246 mg/l, 0.14 mg/l, 15 g/l agar 100 µ/l gentamycin sulfate and 100 µ/l streptomycin sulfate) for 4-5 days. Single colonies were then inoculated into 5 mls of MG/L broth and preincubated overnight in a shaker (180 rpm) at 30°C.

Sterile tomato cotyledon tissue was obtained from 7-8 day old seedlings which had been grown at 24°C, with a 16hr/ 8hr day/night cycle in 100 x 25 mm petri dishes containing MSSV medium: Murashige-Skoog (MS) salts (#1117 Gibco Laboratories, New York), sucrose 30 g/l, Nitsch vitamins (Thomas, B.R., and Pratt, D. Appl. Genet. (1981) 59:215-219), 0.8% agar (pH 6.0). Any tomato species may be used, however, the inbred breeding line UC82B (Department of Vegetable Crops, University of California, Davis) is preferred. The tips and bases of the cotelydons were removed and the center section placed onto a feeder plate for a 24-hour preincubation period in a low light, 40-50 microEinsteins, but no greater than 80 microEinsteins, at 24°C.

Following the preincubation period, the cotyledon explants were then dipped into the agrobacteria suspension (5 x 10^8 bacteria/ml) for approximately 5 minutes, blotted on sterile paper towels and returned to the original tobacco feeder plates. The explants were cocultivated with the agrobacteria for 48 hours on the tobacco feeder plates in low light (see above) at 24°C, then transferred to regeneration medium containing 500 mg/l of carbenicil-

lin disodium salts and at least 100 mg/l of kanamycin sulfate. The regeneration medium is MS sa4s medium with zeatin (2 mg/l), myo-inositol (100 mg/l), sucrose (20 g/l), Nitsche vitamins and 0.8% agar (pH 6.0). After 10 days and subsequently every three weeks, the explants were transferred to fresh regeneration medium containing 500 mg/l of carbenicillin disodium salts and at least 100 mg/l of kanamycin sulfate. Shoots were harvested from 8 weeks onwards and placed on MSSV medium containing carbenicillin (50 mg/l), kanamycin (50 mg/l) and indole-3-butyric acid (1 mg/l). Roots developed in 7-14 days. Plants were then transplanted into soil.

Example 7
Targeting to Apoplast

The hydropath plot of the deduced amino acid sequence of the Leuconostoc dextransucrase shows a highly hydrophilic protein with a hydrophobic N-terminal region. As seen with many other gram-positive secreted proteins, the N-terminal sequence displays characteristics expected of a signal peptide. However, because these sequences are quite different from those seen for eukaryotic secretory proteins, it is possible that a heterologous leader sequence will be desired for efficient translocation of the sucrase enzyme to a desired cell organelle. Various constructs therefore have been designed to optimize targeting of dextransucrase to the fruit apoplast.

PG/Dextransucrase Constructs

Constructs using the full-length dextransucrase gene as well as various 5' deletions of the gene were fused to either the full length polygalacturonase

(PG) leader (314lbp) or the 5' region of the polygalacturonase leader (146 bp). A new dextransucrase construct was made using the polymerase chain reaction (PCR) method. A 1.6 kb XhoI-SalI fragment of pCGN1222 was synthesized and inserted into XhoI digested pCGN2303 to make plasmid pCGN2327. This new construct, pCGN2327, Figure 10, is identical to pCGN2306 but lacks the EcoRV site at the 5' end.

Constructs containing the PG leaders fused to dextransucrase were prepared as follows: The full length PG leader fragment was cloned as a 314 bp XhoI-SalI PCR fragment into the XhoI site of pCGN2327 to give pCGN2330. The 5' PG leader fragment was cloned as a 146 bp XhoI-SalI PCR fragment into the XhoI site of pCGN2327 to give pCGN2331 (Figure 11). Other constructs having either the full-length or a shorter 5' fragment of the PG-leader are fused to the relevant sucrase genes in a like manner. Where it is desirable, for example, to delete 5' bacterial transit peptide regions, portions of the sucrase gene can be deleted to increase activity of the encoded protein. Once the construct is prepared having the desired truncations and/or deletions, the PG leader/dextransucrase regions can then be cloned into the 2A11 and 35S promoter cassettes (pCGN1241 and pCGN2305 respectively) and then into the binary plasmid 7, or other such constructs.

The resulting binary plasmid having the desired constructs therein are used to transform A. tumefaciens and through Agrobacterium-mediated transformation, to obtain transgenic plants.

Example 8
Analysis of Constructs Comprising Transit Peptides

Two PG leader/GUS constructs were made to study the targetting of dextransucrase. The two constructs, pCGN2313 (full length PG leader) and pCGN2325 (5' PG leader) (Fig. 13), were electroporated into tobacco protoplasts and all fractions (media, cell wall debris, and protoplast) were assayed for GUS activity. (Jefferson et al. Plant Molec. Biol. Rep. (1987) 5:387-405.) In initial studies, no GUS activity was detected in any fraction. The GUS gene contains two glycosylation sites. It has been proposed that glycosylation of GUS at one or more of these sites during processing may inactivate the enzyme. Therefore, the electroporation studies were repeated using the antibiotic tunicamycin, which interferes with glycoprotein biosynthesis and prevents glycosylation of GUS. Data from these studies (Table 9) show that when tunicamycin is added, GUS activity is retained and glycosylation appears blocked. Since glycosylation of proteins occurs between the endoplasmic reticulum and the Golgi, the PG leader, full length as well as 5' only, appears to be successfully taking the GUS protein into the secretory pathway. This is also supported by the fact that the GUS activity is reaching the cell wall fraction of the electroporated protoplasts. Western analysis of the protplast fractions will be used to verify these results.

Table 9[1]

GUS Activity In Electroporated Protoplasts

	Without Tunicamycin Cell Wall		With Tunicamycin Cell Wall	
Construct	Protoplast	Debris	Protoplast	Debris
pB1221[2]	2840	471	23450	431
	20250	214	17500	309
pCGN2313	990	104	25700	9590
	940	70	27400	5475
pCGN2325	2660	192	59100	9990
	1180	101	58300	14070

[1]GUS activity is assayed as described by Jefferson et al. (supra). Activity is reported as fluorescence/180 min. reaction/electroporation (fluorescence units at 365 nm excitation/455 nm emission).

[2]pB1221 is a 35S/GUS construct used as a positive control.

Analysis of PG Leader-Dextransucrase Transformed Tissue

Localization of dextransucrase in the apoplast of the tomato fruit can be determined by various methods. One such method is tissue printing of the fruit onto nitrocellulose paper (Cassab and Varner, (1987) J. Cell. Biol., 105:2581-2588). In this procedure a transformed tomato fruit is sliced, washed briefly in distilled water, quickly dried and blotted onto nitrocellulose paper that has previously been soaked in 0.2M $CaCl_2$ and dried. The tissue print is then immediately dried and treated for the detection of dextransucrase using alkaline phosphatase-conjugated second antibody as follows. The nitrocellulose blot is incubated in a blocking buffer consisting of 50 mM Tris-HCl, pH7.5, 200mM NaCl, 0.05% Tween-20 (TTBS) and

1% bovine serum albumin, then incubated in rabbit antidextransucrase antibody in TTBS. Biotinylated goat anti-rabbit IgG is then bound followed by incubation in streptavidin-alkaline phosphatase conjugate. BCIP (5-bromo-4-chloroindoxyl phosphate) and NBT (nitroblue tetrazolium) are used as substrate to visualize the precipitated indoxyl group. Individual cells on the tissue print are then examined under a dissecting microscope for localization of dextransucrase.

Another method used to locate dextransucrase in transformed tomato fruit is immunofluoresence (Vreeland et al., (1984), Planta., 162:506-517) or immunogold (Mauch and Staehelin, (1989), Plant Cell., 1:447-457, Dorel et al., (1989), J. Cell Biol., 108:327-337, Cassab and Varner, (1987), J. Cell Biol., 105:2581-2588) labelling of antigen in thin tissue sections by dextransucrase polyclonal antibody. Transformed tomato fruit are dissected under fixative conditions appropriate for either light or electron microscopic immunocytochemistry (Greenwood and Chrispeels, (1985), Plant Physiol., 79:65-71). Thin sections of the fruit (0.05-1um) then are cut from the fixed tissue, mounted on slides and processed for immunocytochemistry by either immunofluoresence or immunogold labelling techniques. The labelled thin sections are then analyzed under appropriate microscopic conditions for the localization of dextransucrase.

Example 9

Analysis of Transgenic Plants

NPT II Enzyme Assay

Leaves from plants transformed with a construct of interest which regenerated from kanamycin-resistant explant tissue were assayed for expression of the NPT II gene. Tissue was harvested, frozen in

liquid nitrogen, ground frozen in Eppendorf tubes containing equivalent amounts (w/v) of extraction buffer (2.5 mM Tris-HCl, pH 6.8, 0.143 mM β-mercaptoethanol, 0.27 mM leupeptin) then centrifuged for 15 min. The supernatant from leaf homogenates was added to half volume aliquots of reaction buffers A and B. Reaction buffer A contained 67 mM Tris-maleate, 42 mM $MgCl_2$, 400 mM NH_4Cl, 1.7 mM dithiothreitol, 0.4 mg/ml kanamycin sulfate. Reaction buffer B was identical to buffer A, except that kanamycin was omitted. [γ-^{32}P] ATP solution (0.75 mM ATP dissolved in reaction buffer B and 1.0 uCi [γ-^{32}P] ATP/sample) was added to each reaction mixture. The samples were incubated at 32°C for 30 min. Using a slot blotting apparatus, reaction samples were blotted onto three layers of Whatman P81 ion exchange papers placed on top of one Whatman 3mm paper. All three P81 papers were washed twice with water at 25°C for 4 min each. The blots were incubated for 45 min at 65°C in proteinase K solution (1.0 mg/ml proteinase K in 1% SDS), then washed once at 80°C for 4 min followed by a wash at room temperature. After drying, blots were exposed to X-ray film with an intensifying screen.

Northern Analysis

Out of 19 transformed plants containing the levansucrase construct, and 2 transformed plants containing the 2A11/dextransucrase construct, and 42 transformed plants containing the 355/dextransucrase construct, 14, 41 and 35 plants, respectively, show expression the kan gene.

Isolation of Poly(A)$^+$ RNA

Leaves and fruit (green as well as ripe) from transformed tomato plants were harvested and frozen in liquid nitrogen. Frozen leaves were ground in a mortar and pestle in liquid nitrogen and total RNA was extrac-

ted from the resulting powder by homogenization in Tris-HCl, SDS buffer and lithium chloride precipitation as described by Crouch et al., J. Mol. Appl. Gen. (1983) 2:273-283. Frozen ripe fruit was ground to a powder in a mortar and pestle in liquid nitrogen and homogenized with a Brinkmann polytron continuously for 2 min on ice. Total RNA is isolated as described by Colbert et al., Proc. Natl. Acad. Sci. (1983) 80:2248-2252 using the extraction buffer described by Facciotti et al., Bio/Tech (1985) 3:241-246 and pelleted through a high-density CsCl solution on a Beckman type 60Ti rotor as described by Turpen and Griffith, BioTechniques (1986) 4:11-15. Poly(A)$^+$ RNA is isolated as described by Maniatis et al. (1982). Contaminating sucrases are removed from the samples by washing with Sigmacell ^{R} type 50 in 20 mM Tris-HCl, pH 7.5, 0.5 M NaCl, 1 mM EDTA, 0.1% SDS.

RNA Probe Preparation

A ^{32}P-labeled RNA probe complementary to the gene of interest is synthesized using a Riboprobe kit (Promega Biotech), according to the manufacturers instructions, and labeling with α^{32}P-rUTP, according to the manufacturer's instructions. The probe to the levansucrase gene was synthesized from the T7 promoter of the plasmid pCGN499 (see supra).

Analysis

mRNA isolated from leaves of transformed plants is denatured and electrophoresed as described by Shewmaker et al., Virology (1985) 140:281-288. The gels were blotted overnight against 20x SSC (3 M sodium chloride, 0.3 M sodium citrate, pH 7.0) the nitrocellulose filters baked at 80°C for 2 hrs and then hybridized at 50°C in the hybridization buffer suggested by the Riboprobe manufacturer containing 0.5 mg/ml denatured yeast tRNA, 0.5 mg/ml denatured salmon sperm DNA, and 0.25 mg/ml polyanethole-

sulfonic acid. Blots are then washed at 55°C for 30 min and twice at 65°C for 30 min each in buffer as described by the Riboprobe manufacturer. Contaminating ribosomal RNA is removed from the blots by washing them with 1 ug/ml ribonuclease in 2x SSC at 25°C.

As shown in Figure 12, two of the transformed plants analyzed hybridized to the levansucrase probe, although only one of the constructs was producing a full length transcript. Analysis of mRNA from ripe fruit from the transformants gave the same results.

Analysis of Tomato Fruit for Sucrase Activity

Protein is isolated from green fruit from a transformed plant, which appears from Northern analysis to be producing a full length transcript of the gene of interest, and from green fruit of nontransformed Calgrande plants as a negative control. The method used is as follows. Fruit is frozen then allowed to thaw in 1.5 M NaCl, 0.1 M sodium citrate, pH 6 buffer, ground in a mortar and pestle, then homogenized using a Polytron for 3 min at 4°C. Large debris is removed by centrifugation twice at 5500 x g. The supernatant is then brought to 95% $(NH_4)_2SO_4$. After stirring for 1 hr at 4°C the samples are centrifuged and the protein precipitate is resuspended in 10 ml of 0.05 M KPO_4 buffer, pH 5. The samples are then dialyzed overnight against 0.05 M KPO_4, pH 6, centrifuged to remove any precipitants and the supernatant analyzed for levansucrase activity.

Levansucrase was isolated from Bacillus subtilis strain QB118 as described by Dedonder, Methods in Enzymol. (1966) 8:500-505. Briefly, the supernatant from a one-liter 14-hr culture grown in EcLB (5 g/l yeast extract, 10 g/l tryptone, 5 g/l NaCl) containing 10 g/l glucose at 30°C was isolated and pH adjusted to 4.2 by addition of acetic acid. Cold ethanol (2-4°C) was gradually added to 48% and stirred continuously at 4°C for 20 min then centrifuged to pellet the proteins. The pellet was dissolved

in 30 ml of 0.05 M KPO_4 and dialyzed overnight against 50% $(NH_4)_2SO_4$, pH 5, at 4°C overnight. The sample was then centrifuged and the supernatant dialyzed against 65% $(NH_4)_2SO_4$, pH 5, at 4°C overnight. Again the sample was centrifuged and the supernatant was dialyzed against 95% $(NH_4)_2SO_4$, pH 5, at 4°C overnight. The sample was then centrifuged to pellet the proteins and the pellet was resuspended in 5 ml of 0.05 M KPO_4, pH 6, and dialyzed against 0.05 M KPO_4, pH 6, at 4°C overnight.

Dextransucrase process a radioactive polymer only from sucrose labelled in the glucose molecule of sucrase, while levansucrase produces a radioactive polymer only from sucrose labeled in the fructose molecule of sucrose. The protocol for listing labelled [glucose-^{14}C] as required in analysis of dextransucrase is discussed above. 50 µl of the supernatant obtained as above was added to reaction tubes containing 0.1µCi of dried sucrose [fructose-^{14}C] and 50µl of reaction buffer (0.05M KPO_4, pH6, 0.01M sucrose, 20mg/ml levan (10,000 MW). The reaction tubes were incubated at 30°C and sampled at various times for the incorporation of ^{14}C into polymer by spotting aliquots of each reaction onto Whatman No. 1 filter paper. The filters were then washed in methanol as described supra for analysis of dextransucrase.

The above results demonstrate that plant species can be transformed efficiently with expression cassettes which provide for expression of a heterologous gene for a carbohydrate polymer synthesizing enzyme such as a bacterial dextran or levansucrase. As evidenced by the above disclosure, plant species are provided which express a bacterial levansucrase or dextransucrase gene and as a result may have modified solids content without adverse effects on the normal production capabilities of the plant.

All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this

invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the appended claims.

WHAT IS CLAIMED IS:

1. A method for modifying the soluble solids composition of a plant cell, said method comprising:

growing plant cells comprising at least one copy of a foreign gene encoding an enzyme capable of polymerizing a carbohydrate comprising at least one hexose moiety to form a polymer under conditions whereby said gene is expressed and the soluble solids composition of said plant cells is modified.

2. The method according to Claim 1, wherein said carbohydrate is sucrose.

3. The method according to Claim 2, wherein said enzyme is dextransucrase or levansucrase.

4. A method for modifying the soluble solids composition of fruit, said method comprising:

growing a plant comprising cells containing at least one copy of an expression cassette comprising, in the 5'-3' direction of transcription, a transcriptional and translational initiation region active in a fruit cell, a DNA sequence encoding an enzyme capable of polymerizing a carbohydrate comprising at least one hexose moiety to form a polymer, wherein expression of said DNA sequence is regulated by said initiation region, and translational and transcriptional termination regions functional in a plant cell, under conditions whereby said DNA sequence is expressed and the soluble solids composition of said fruit is modified.

5. A plant cell capable of producing sucrase at an enhanced level, said cell comprising:

at least one copy of a sucrase gene, wherein said cell was obtained by inserting into said cell, or a parent of said cell, an expression cassette, prepared in vitro,

comprising a sucrase gene under the transcriptional and translational control of transcriptional and translational initiation and termination regulatory regions functional in said cell.

6. The plant cell according to Claim 5, wherein said sucrase gene is a levansucrase gene or a dextransucrase gene.

7. The plant cell according to Claim 5, wherein said plant cell is a fruit cell or a seed cell.

8. The plant cell according to Claim 6, wherein said plant cell is a tomato plant cell.

9. A plant producing increased levels of sucrase, said plant consisting essentially of:

cells obtained by inserting into said cells, or parents of said cells, an expression cassette, prepared in vitro, comprising a sucrase gene under the transcriptional and translational control of transcriptional and translational initiation and termination regulatory regions functional in said cells, whereby said cells produce increased levels of sucrase.

10. The plant according to Claim 9, wherein said sucrase gene is a levansucrase gene or a dextransucrase gene.

11. A tomato plant comprising cells producing levansucrase or dextransucrase.

12. The plant according to Claim 11, wherein said cells are fruit cells.

13. An expression cassette comprising a transcriptional and translational region functional in a

plant cell, a DNA sequence encoding levansucrase, and a transcriptional termination regulatory functional in a plant cell.

14. A plant cell comprising an expression cassette of Claim 13.

15. A DNA fragment at least substantially homologous with a DNA sequence encoding dextransucrase obtainable from L. mesenteroides.

16. A DNA fragment comprising L. mesenteroides dextransucrase structural gene.

17. A DNA fragment comprising the DNA sequence shown in Figure 4.

18. The DNA fragment according to any one of Claims 15, 16, or 17, wherein said fragment is joined to a heterologous DNA sequence.

19. A DNA construct comprising a DNA sequence encoding dextransucrase obtainable from L. mesenteroides joined to a heterologous DNA sequence.

20. The DNA construct according to Claim 19, wherein said heterologous DNA sequence is a transcriptional and translational initiation region functional in a plant cell.

21. The DNA construct according to Claim 20, wherein said transcriptional and translational initiation region is active at or immediately after anthesis and during fruit ripening.

22. The DNA construct according to Claim 21, wherein said initiation region is the 2A11 initiation region.

23. An expression cassette comprising: in the 5'-3' direction of transcription, a transcriptional and translational region derived from 2A11; a first DNA sequence encoding dextransucrase or levansucrase, and a transcriptional and translational termination regulatory region functional in a plant cell.

24. An expression cassette according to Claim 23, further comprising:

a second DNA sequence encoding a transit peptide joined in reading frame at the 5' terminus of said first DNA sequence, wherein said second DNA sequence is derived from a gene encoding an apoplast polypeptide.

25. An expression cassette according to Claim 24, wherein said apoplast polypeptide is tomato poly-galacturonidase.

26. A cell comprising a DNA construct according to Claim 25.

27. The cell according to Claim 26, wherein said cell is a plant cell.

28. The cell according to Claim 27, wherein said plant cell is in vivo.

29. The cell according to Claim 28, wherein said plant is tomato.

30. The cell according to Claim 29 comprising:

a DNA construct according to any one of Claims 20, 21, 22, 23, 24, and 25.

31. A polypeptide comprising:

a sucrase joined to a plant apoplast transit peptide sequence.

32. The polypeptide according to Claim 31, wherein said sucrase is dextransucrase or levansucrase.

33. The polypeptide according to Claim 32, wherein said dextransucrase is L. mesenteroides dextransucrase.

34. The polypeptide according to Claim 31, wherein said plant apoplast transit peptide sequence is derived from tomato polygalacturonase.

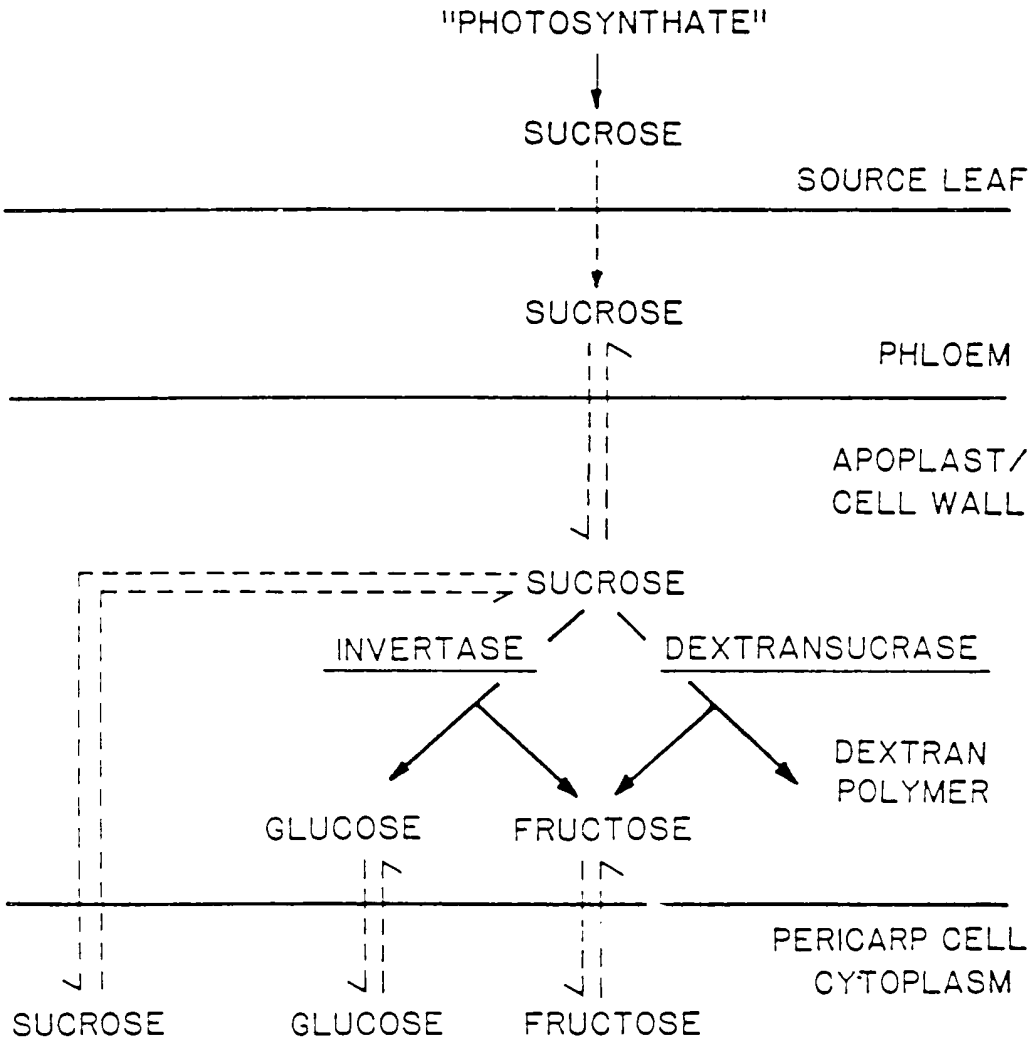

"PHOTOSYNTHATE"
SUCROSE
SOURCE LEAF
SUCROSE
PHLOEM
APOPLAST/
CELL WALL
SUCROSE
INVERTASE
DEXTRANSUCRASE
DEXTRAN
POLYMER
GLUCOSE
FRUCTOSE
PERICARP CELL
CYTOPLASM
SUCROSE
GLUCOSE
FRUCTOSE

FIG. 1

FIG. 2-1

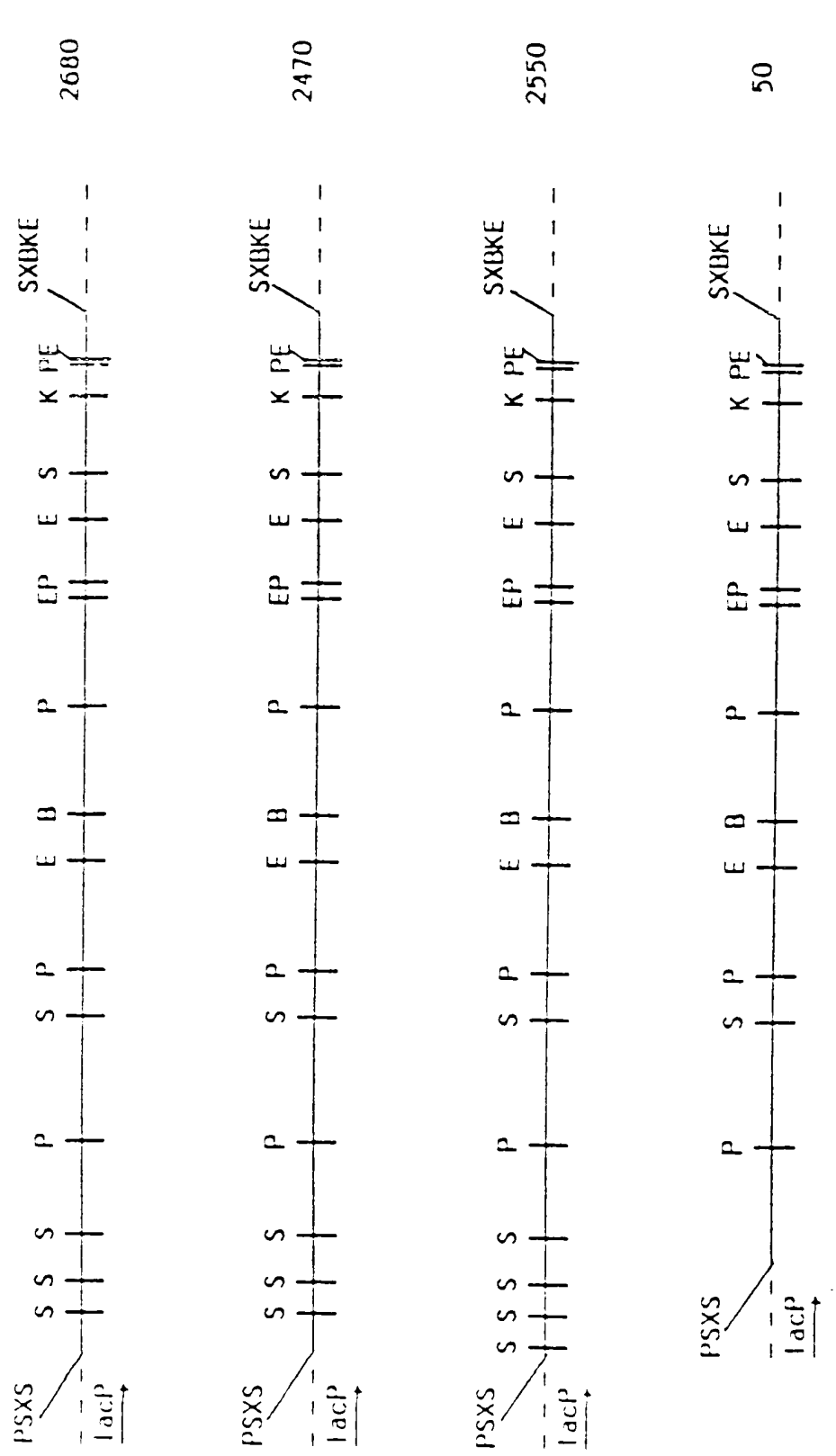
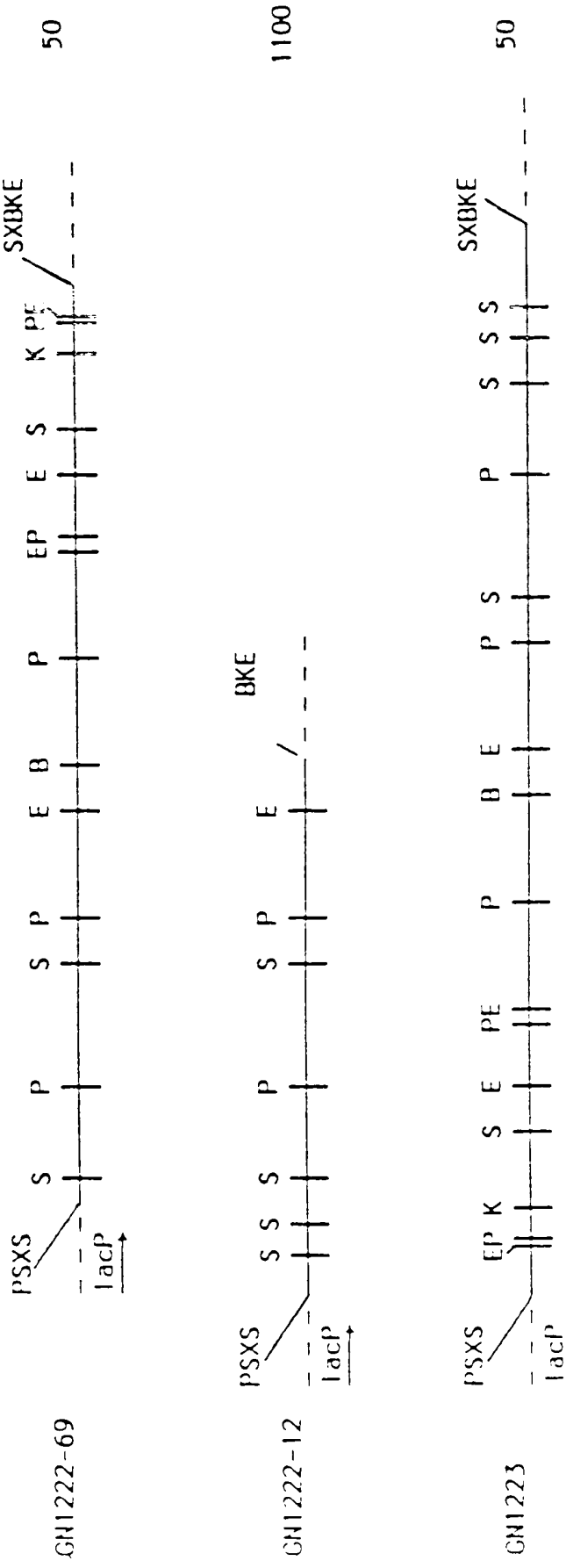

ACTIVITY
2680
2470
2550
50
SXBKE
PSXS
lacP
S S S P P E B P EP E S K PE
N1222
N1222-16
N1222-18
N1222-1

FIG. 2-2


50
1100
50
SXBKE
PSXS
S P E B P EP E S K PE
lacP
GN1222-69
BKE
PSXS
S S S P S P E
lacP
GN1222-12
SXBKE
PSXS
EP K S E PE P B E P S P S S S
lacP
GN1223

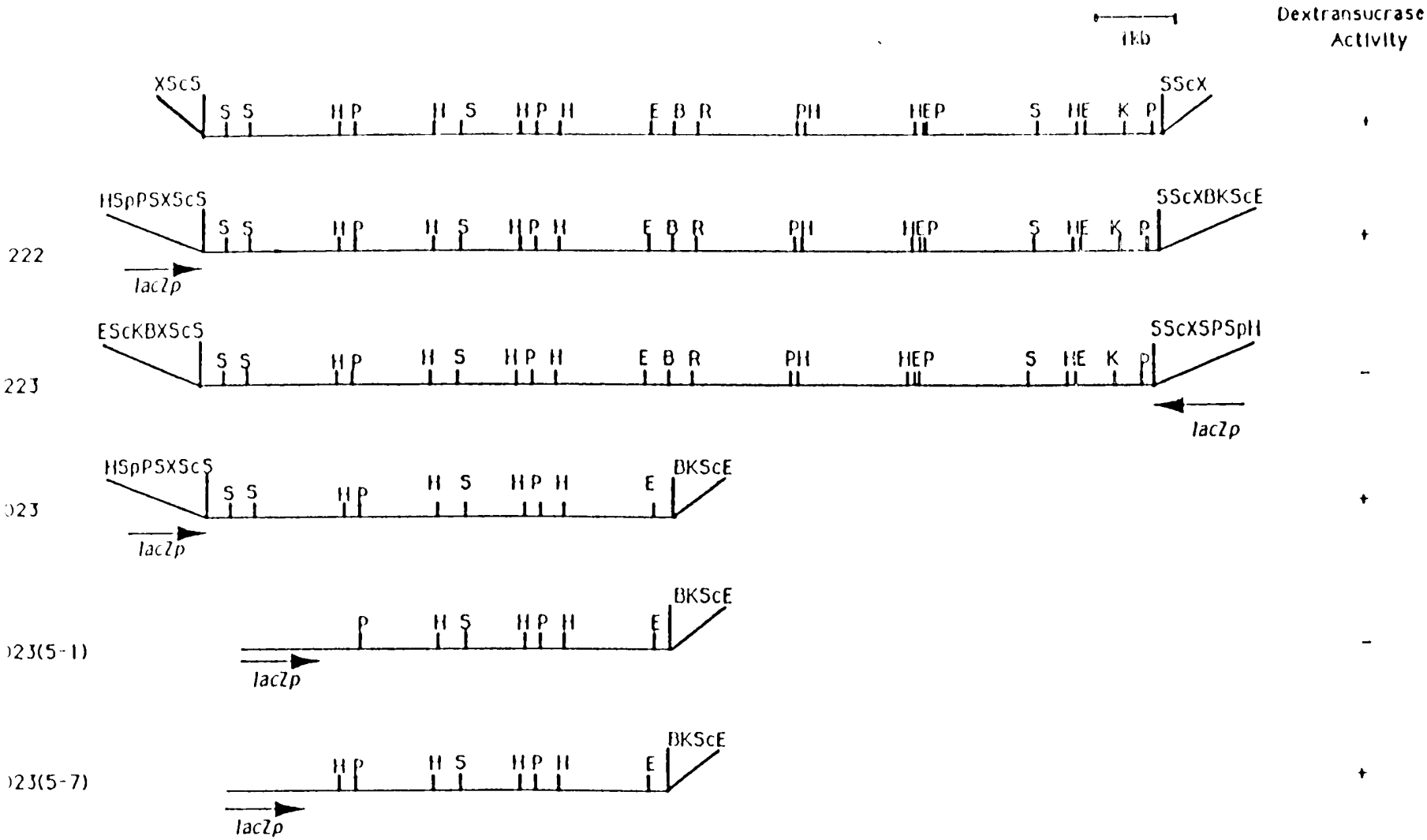

1kb
Dextransucrase Activity
XScS
SScX
HSpPSXScS
SScXBKScE
222
lacZp
EScKBXScS
SScXSPSpH
223
lacZp
HSpPSXScS
BKScE
023
lacZp
BKScE
023(5-1)
lacZp
BKScE
023(5-7)
lacZp

FIG. 3

G. 4-1

```
AATAAAATTAATGACTAATTAATAATAATTAACATTGTTAACTTATCACTAGATAAGTTAACAATGTTTTT     71
                                    ---------------->    <----------------
TTATAAAAAAATTAATGTTTATAATATTAATATAATTCTCTGATTGACTGTTTTTACAATGTTTTTTAAAA    142
ATATGTCGATTGTGTATTGAAATTTTTACTGTTTTTACAAATAAATTAAGTTATAATTTTTCAATGTAATA    213
                      -35                       -10
GAAGAGAATATTATAAGGAGAAAAATTT ATG CCA TTT ACA GAA AAA GTA ATG CGG AAA AAG    273
                      SD          M   P   F   T   E   K   V   M   R   K   K
CTT TAT AAA GTT GGG AAA AGT TGG GTA GTT GGT GGG GTT TGT GCT TTT GCA TTA    327
L   Y   K   V   G   K   S   W   V   V   G   G   V   C   A   F   A   L
ACC GCC TCA TTT GCT TTA GCA ACA CCA AGT GTT TTA GGA GAC AGT AGT GTA CCT    381
T   A   S   F   A   L   A   T   P   S   V   L   G   D   S   S   V   P
GAT GTG AGT GCG AAT AAC GTT CAA TCT GCT TCA GAT AAT ACA ACG GAT ACG CAG    435
D   V   S   A   N   N   V   Q   S   A   S   D   N   T   T   D   T   Q
CAG AAC ACT ACG GTT ACC GAA GAA AAT GAT AAA GTA CAG TCT GCA GCT ACT AAT    489
Q   N   T   T   V   T   E   E   N   D   K   V   Q   S   A   A   T   N
GAC AAT GTA ACA ACA GCT GCA AGC GAC ACA ACA CAA TCT GCT GAT AAT AAT GTG    543
D   N   V   T   T   A   A   S   D   T   T   Q   S   A   D   N   N   V
ACA GAA AAA CAG TCA GAT GAT CAT GCA CTT GAT AAT GAA AAA GTC GAT AAC AAA    597
T   E   K   Q   S   D   D   H   A   L   D   N   E   K   V   D   N   K
CAA GAT GAA GTC GCT CAA ACC AAT GTT ACT AGC AAA AAT GAG GAA TCA GCA GTT    651
Q   D   E   V   A   Q   T   N   V   T   S   K   N   E   E   S   A   V
GCT TCA ACT GAC ACT GAT CCT GCT GAA ACG ACA ACT GAC GAA ACA CAA CAA GTT    705
A   S   T   D   T   D   P   A   E   T   T   T   D   E   T   Q   Q   V
AGC GGC AAG TAC GTT GAA AAA GAC GGT AGT TGG TAT TAT TAT TTT GAT GAT GGC    759
S   G   K   Y   V   E   K   D   G   S   W   Y   Y   Y   F   D   D   G
AAA AAT GCT AAA GGT TTA TCA ACG ATA GAC AAC AAT ATT CAA TAT TTT TAC GAG    813
K   N   A   K   G   L   S   T   I   D   N   N   I   Q   Y   F   Y   E
```

G. 4-2

```
AGT GGT AAA CAA GCC AAA GGA CAG TAT GTC ACA ATT GAT AAT CAA ACA TAT TAT   867
S   G   K   Q   A   K   G   Q   Y   V   T   I   D   N   Q   T   Y   Y
TTT GAT AAG GGC TCA GGT GAT GAG TTA ACT GGT CTG CAA AGC ATT GAT GGG AAC   921
F   D   K   G   S   G   D   E   L   T   G   L   Q   S   I   D   G   N
ATA GTT GCT TTT AAC GAT GAA GGG CAA CAA ATT TTT AAT CAA TAT TAC CAA TCT   975
I   V   A   F   N   D   E   G   Q   Q   I   F   N   Q   Y   Y   Q   S
GAA AAT GGT ACA ACA TAC TAC TTT GAT GAT AAA GGA CAC GCT GCT ACC GGT ATT  1029
E   N   G   T   T   Y   Y   F   D   D   K   G   H   A   A   T   G   I
AAG AAT ATC GAG GGC AAA AAT TAT TAT TTT GAT AAT CTT GGG CAA CTA AAA AAA  1083
K   N   I   E   G   K   N   Y   Y   F   D   N   L   G   Q   L   K   K
GGC TTC TCT GGT GTG ATT GAT GGT CAA ATA ATG ACA TTT GAT CAG GAA ACA GGG  1137
G   F   S   G   V   I   D   G   Q   I   M   T   F   D   Q   E   T   G
CAA GAA GTT TCT AAC ACA ACT TCT GAA ATA AAA GAA GGT TTG ACG ACT CAA AAC  1191
Q   E   V   S   N   T   T   S   E   I   K   E   G   L   T   T   Q   N
ACG GAT TAT AGC GAA CAT AAT GCA GCC CAC GGT ACG GAT GCT GAG GAC TTT GAA  1245
T   D   Y   S   E   H   N   A   A   H   G   T   D   A   E   D   F   E
AAT ATT GAC GGC TAT TTA ACA GCT AGT TCA TGG TAT CGT CCA ACA GGT ATT TTA  1299
N   I   D   G   Y   L   T   A   S   S   W   Y   R   P   T   G   I   L
CGT AAC GGA ACA GAC TGG GAA CCT TCT ACA GAT ACA GAT TTC AGA CCA ATA TTG  1353
R   N   G   T   D   W   E   P   S   T   D   T   D   F   R   P   I   L
TCA GTG TGG TGG CCA GAT AAG AAC ACC CAG GTC AAT TAT TTA AAT TAC ATG GCT  1407
S   V   W   W   P   D   K   N   T   Q   V   N   Y   L   N   Y   M   A
GAT TTA GGG TTT ATC AGT AAT GCG GAC AGT TTT GAA ACT GGG GAT AGC CAA AGC  1461
D   L   G   F   I   S   N   A   D   S   F   E   T   G   D   S   Q   S
TTA TTA AAT GAA GCA AGT AAC TAT GTT CAA AAA TCA ATT GAA ATG AAA ATT AGT  1515
L   L   N   E   A   S   N   Y   V   Q   K   S   I   E   M   K   I   S
GCG CAA CAA AGT ACA GAG TGG TTA AAG GAT GCA ATG GCG GCC TTC ATT GTC GCG  1569
A   Q   Q   S   T   E   W   L   K   D   A   M   A   A   F   I   V   A
CAA CCA CAG TGG AAT GAA ACT AGT GAA GAT ATG AGC AAT GAC CAT TTA CAA AAT  1623
Q   P   Q   W   N   E   T   S   E   D   M   S   N   D   H   L   Q   N
GGC GCA TTA ACT TAT GTC AAC AGT CCA CTG ACA CCT GAC GCT AAT TCA AAC TTT  1677
G   A   L   T   Y   V   N   S   P   L   T   P   D   A   N   S   N   F
```

Fig. 4-3

```
AGA CTA CTT AAT CGG ACA CCA ACA AAC CAG ACT GGT GAA CAA GCC TAT AAT TTA  1731
R   L   L   N   R   T   P   T   N   Q   T   G   E   Q   A   Y   N   L
GAT AAT TCA AAA GGT GGT TTT GAA TTG TTG TTA GCC AAT GAC GTT GAT AAT TCA  1785
D   N   S   K   G   G   F   E   L   L   L   A   N   D   V   D   N   S
AAC CCT GTA GTA CAA GCA GAA CAA TTG AAT TGG TTA TAT TAT TTA ATG AAT TTT  1839
N   P   V   V   Q   A   E   Q   L   N   W   L   Y   Y   L   M   N   F
GGT ACG ATT ACG GCC AAC GAC GCG GAT GCT AAT TTT GAT GGT ATT CGT GTA GAT  1893
G   T   I   T   A   N   D   A   D   A   N   F   D   G   I   R   V   D
GCA GTC GAC AAT GTG GAT GCT GAT TTG TTA CAA ATT GCT GCC GAT TAT TTC AAA  1947
A   V   D   N   V   D   A   D   L   L   Q   I   A   A   D   Y   F   K
CTA GCT TAC GGT GTT GAT CAA AAT GAT GCT ACT GCT AAT CAG CAT CTT TCA ATT  2001
L   A   Y   G   V   D   Q   N   D   A   T   A   N   Q   H   L   S   I
TTG GAA GAT TGG AGT CAC AAT GAT CCT TTG TAT GTA ACA GAT CAA GGA AGC AAT  2055
L   E   D   W   S   H   N   D   P   L   Y   V   T   D   Q   G   S   N
CAA TTA ACC ATG GAT GAT TAT GTG CAC ACA CAA TTA ATC TGG TCT CTA ACA AAA  2109
Q   L   T   M   D   D   Y   V   H   T   Q   L   I   W   S   L   T   K
TCA TCT GAC ATA CGA GGT ACA ATG CAG CGC TTC GTG GAT TAT TAT ATG GTG GAT  2163
S   S   D   I   R   G   T   M   Q   R   F   V   D   Y   Y   M   V   D
CGA TCT AAT GAT AGT ACA GAA AAC GAA GCC ATT CCT AAT TAC AGC TTT GTA CGT  2217
R   S   N   D   S   T   E   N   E   A   I   P   N   Y   S   F   V   R
GCA CAC GAC AGC GAA GTG CAA ACG GTT ATT GCC CAA ATT GTT TCC GAT TTG TAT  2271
A   H   D   S   E   V   Q   T   V   I   A   Q   I   V   S   D   L   Y
CCT GAT GTT GAA AAT AGT TTA GCA CCA ACA ACA GAA CAA TTG GCA GCT GCT TTC  2325
P   D   V   E   N   S   L   A   P   T   T   E   Q   L   A   A   A   F
AAA GTA TAC AAT GAA GAT GAA AAA TTA GCA GAC AAA AAG TAC ACA CAA TAT AAT  2379
K   V   Y   N   E   D   E   K   L   A   D   K   K   Y   T   Q   Y   N
ATG GCT AGT GCT TAT GCG ATG TTG CTA ACC AAT AAG GAT ACT GTT CCT CGT GTC  2433
M   A   S   A   Y   A   M   L   L   T   N   K   D   T   V   P   R   V
TAT TAT GGC GAT TTA TAT ACA GAT GAT GGT CAA TAT ATG GCA ACA AAG TCA CCA  2487
Y   Y   G   D   L   Y   T   D   D   G   Q   Y   M   A   T   K   S   P
```

FIG.4-4

```
TAC TAT GAT GCG ATT AAC ACT TTG CTA AAG GCT AGA GTT CAG TAT GTT GCT GGT  2541
Y   Y   D   A   I   N   T   L   L   K   A   R   V   Q   Y   V   A   G
GGC CAA TCG ATG TCC GTT GAT AGT AAT GAC GTG TTA ACA AGT GTT CGC TAT GGT  2595
G   Q   S   M   S   V   D   S   N   D   V   L   T   S   V   R   Y   G
AAA GAT GCC ATG ACA GCT TCT GAC ACT GGA ACA TCT GAG ACG CGT ACT GAA GGT  2649
K   D   A   M   T   A   S   D   T   G   T   S   E   T   R   T   E   G
ATT GGA GTC ATC GTC AGC AAT AAC GCG GAG CTA CAA TTA GAG GAT GGG CAT ACT  2703
I   G   V   I   V   S   N   N   A   E   L   Q   L   E   D   G   H   T
GTC ACA TTG CAT ATG GGG GCA GCT CAT AAG AAC CAA GCT TAT CGT GCT TTG TTA  2757
V   T   L   H   M   G   A   A   H   K   N   Q   A   Y   R   A   L   L
TCA ACA ACT GCA GAT GGA TTA GCT TAT TAT GAT ACT GAT GAA AAT GCA CCT GTG  2811
S   T   T   A   D   G   L   A   Y   Y   D   T   D   E   N   A   P   V
GCG TAC ACA GAT GCT AAC GGC GAT TTG ATT TTT ACG AAT GAA TCA ATT TAT GGT  2865
A   Y   T   D   A   N   G   D   L   I   F   T   N   E   S   I   Y   G
GTA CAA AAT CCA CAA GTT TCT GGT TAC TTG GCA GTT TGG GTT CCG GTA GGT GCG  2919
V   Q   N   P   Q   V   S   G   Y   L   A   V   W   V   P   V   G   A
CAA CAA GAT CAA GAT GCA CGA ACG GCC TCT GAT ACA ACA ACA AAC ACG AGT GAT  2973
Q   Q   D   Q   D   A   R   T   A   S   D   T   T   T   N   T   S   D
AAA GTG TTC CAT TCA AAC GCT GCT CTT GAT TCT CAA GTC ATC TAC GAA GGT TTC  3027
K   V   F   H   S   N   A   A   L   D   S   Q   V   I   Y   E   G   F
TCA AAC TTC CAA GCA TTT GCT ACA GAC AGC AGT GAA TAT ACA AAC GTA GTC ATC  3081
S   N   F   Q   A   F   A   T   D   S   S   E   Y   T   N   V   V   I
GCT CAG AAT GCG GAC CAA TTT AAG CAA TGG GGT GTG ACA AGC TTC CAA TTG GCA  3135
A   Q   N   A   D   Q   F   K   Q   W   G   V   T   S   F   Q   L   A
CCA CAA TAT CGT TCA AGT ACA GAT ACA AGT TTC TTG GAT TCA ATT ATT CAA AAC  3189
P   Q   Y   R   S   S   T   D   T   S   F   L   D   S   I   I   Q   N
GGG TAT GCA TTC ACG GAT CGT TAT GAC TTA GGT TAT GGC ACA CCG ACA AAA TAT  3243
G   Y   A   F   T   D   R   Y   D   L   G   Y   G   T   P   T   K   Y
GGA ACT GCT GAT CAG TTG CGC GAT GCT ATT AAA GCC TTA CAT GCT AGC GGT ATT  3297
G   T   A   D   Q   L   R   D   A   I   K   A   L   H   A   S   G   I
```

G.4-5

```
CAA GCC ATT GCC GAT TGG GTG CCG GAC CAA ATT TAT AAT TTG CCA GAG CAA GAA  3351
Q   A   I   A   D   W   V   P   D   Q   I   Y   N   L   P   E   Q   E
TTA GCT ACT GTC ACA AGA ACA AAT TCA TTT GGA GAT GAC GAT ACA GAT TCT GAT  3405
L   A   T   V   T   R   T   N   S   F   G   D   D   D   T   D   S   D
ATT GAC AAT GCC TTA TAT GTT GTA CAA AGT CGT GGG GGT GGT CAA TAT CAA GAG  3459
I   D   N   A   L   Y   V   V   Q   S   R   G   G   G   Q   Y   Q   E
ATG TAT GGT GGT GCC TTC TTA GAA GAG TTA CAG GCA CTC TAT CCA TCC CTA TTT  3513
M   Y   G   G   A   F   L   E   E   L   Q   A   L   Y   P   S   L   F
AAA GTG AAT CAA ATC TCA ACT GGC GTT CCA ATT GAT GGC AGT GTA AAG ATT ACT  3567
K   V   N   Q   I   S   T   G   V   P   I   D   G   S   V   K   I   T
GAG TGG GCG GCT AAG TAC TTC AAT GGC TCT AAC ATC CAA GGT AAA GGT GCT GGA  3621
E   W   A   A   K   Y   F   N   G   S   N   I   Q   G   K   G   A   G
TAC GTA TTG AAA GAT ATG GGT TCT AAT AAG TAC TTT AAG GTC GTT TCG AAC ACT  3675
Y   V   L   K   D   M   G   S   N   K   Y   F   K   V   V   S   N   T
GAG GAT GGT GAC TAC TTA CCA AAA CAG TTA ACT AAT GAT CTG TCA GAA ACT GGC  3729
E   D   G   D   Y   L   P   K   Q   L   T   N   D   L   S   E   T   G
TTT ACA CAC GAT GAT AAA GGA ATC ATC TAT TAT ACA TTA AGT GGT TAT CGT GCC  3783
F   T   H   D   D   K   G   I   I   Y   Y   T   L   S   G   Y   R   A
CAA AAT GCA TTT ATT CAA GAT GAT GAT AAT AAC TAT TAC TAT TTT GAT AAA ACA  3837
Q   N   A   F   I   Q   D   D   D   N   N   Y   Y   Y   F   D   K   T
GGT CAT TTA GTA ACA GGT TTG CAA AAG ATT AAT AAC CAT ACC TAC TTC TTC TTA  3891
G   H   L   V   T   G   L   Q   K   I   N   N   H   T   Y   F   F   L
CCT AAT GGT ATC GAA CTG GTC AAG AGC TTC TTA CAA AAC GAA GAT GGT ACA ATT  3945
P   N   G   I   E   L   V   K   S   F   L   Q   N   E   D   G   T   I
GTT TAT TTC GAT AAG AAA GGT CAT CAA GTT TTT GAT CAA TAT ATA ACT GAT CAA  3999
V   Y   F   D   K   K   G   H   Q   V   F   D   Q   Y   I   T   D   Q
AAT GGA AAT GCG TAT TAC TTT GAT GAT GCT GGT GTA ATG CTT AAA TCA GGG CTT  4053
N   G   N   A   Y   Y   F   D   D   A   G   V   M   L   K   S   G   L
GCA ACG ATT GAT GGA CAT CAA CAG TAT TTT GAT CAA AAT GGT GTC CAG GTT AAG  4107
A   T   I   D   G   H   Q   Q   Y   F   D   Q   N   G   V   Q   V   K
```

G.4-6

```
GAT AAG TTT GTG ATT GGC ACT GAT GGT TAT AAG TAT TAC TTT GAA CCA GGT AGT  4161
D   K   F   V   I   G   T   D   G   Y   K   Y   Y   F   E   P   G   S
GGT AAC TTA GCT ATC CTA CGT TAT GTG CAA AAT AGT AAG AAT CAA TGG TTC TAT  4215
G   N   L   A   I   L   R   Y   V   Q   N   S   K   N   Q   W   F   Y
TTT GAT GGT AAT GGC CAT GCT GTC ACT GGT TTC CAA ACA ATT AAT GGT AAA AAA  4269
F   D   G   N   G   H   A   V   T   G   F   Q   T   I   N   G   K   K
CAA TAT TTC TAT AAT GAT GGT CAT CAA AGT AAA GGT GAA TTC ATT GAT GCA GAC  4323
Q   Y   F   Y   N   D   G   H   Q   S   K   G   E   F   I   D   A   D
GGG GAT ACT TTC TAT ACG AGT GCC ACT GAT GGT CGC CTA GTA ACT GGT GTT CAG  4377
G   D   T   F   Y   T   S   A   T   D   G   R   L   V   T   G   V   Q
AAG ATT AAT GGT ATT ACC TAT GCT TTT GAT AAC ACA GGA AAT TTG ATC ACA AAT  4431
K   I   N   G   I   T   Y   A   F   D   N   T   G   N   L   I   T   N
CAG TAT TAT CAA TTA GCA GAT GGT AAA TAT ATG TTG TTA GAT GAT AGT GGT CGT  4485
Q   Y   Y   Q   L   A   D   G   K   Y   M   L   L   D   D   S   G   R
GCG AAA ACA GGG TTT GTA TTG CAA GAT GGT GTA CTA AGA TAC TTC GAT CAA AAC  4539
A   K   T   G   F   V   L   Q   D   G   V   L   R   Y   F   D   Q   N
GGT GAG CAA GTG AAA GAT GCT ATC ATT GTG GAT CCA GAT ACT AAC TTG AGT TAT  4593
G   E   Q   V   K   D   A   I   I   V   D   P   D   T   N   L   S   Y
TAT TTC AAT GCA ACA CAA GGT GTC GCT GTA AAA AAT GAT TAT TTC GAG TAT CAA  4647
Y   F   N   A   T   Q   G   V   A   V   K   N   D   Y   F   E   Y   Q
GGT AAT TGG TAT TTA ACA GAT GCT AAT TAT CAA CTT ATC AAA GGT TTT AAA GCA  4701
G   N   W   Y   L   T   D   A   N   Y   Q   L   I   K   G   F   K   A
GTT GAC GAC AGC TTA CAA CAT TTT GAT GAA GTC ACT GGT GTA CAA ACA AAA GAT  4755
V   D   D   S   L   Q   H   F   D   E   V   T   G   V   Q   T   K   D
AGT GCT TTA ATA AGT GCT CAG GGT AAG GTT TAC CAA TTT GAT AAT AAT GGA AAT  4809
S   A   L   I   S   A   Q   G   K   V   Y   Q   F   D   N   N   G   N
GCT GTG TCA GCA TAA GCTTTCTGTATATATAGTGAAAAGCCAAAGGTCCTTTTGACCTTTGGCTTT  4875
A   V   S   A   END                        ---------->     <----------
TTGATTTAATTGTCTATTTGAATTTCCTTGATGATATCGGCTAATTTAATTTTATTCATTTCTTTTTCGGC  4946
--
```

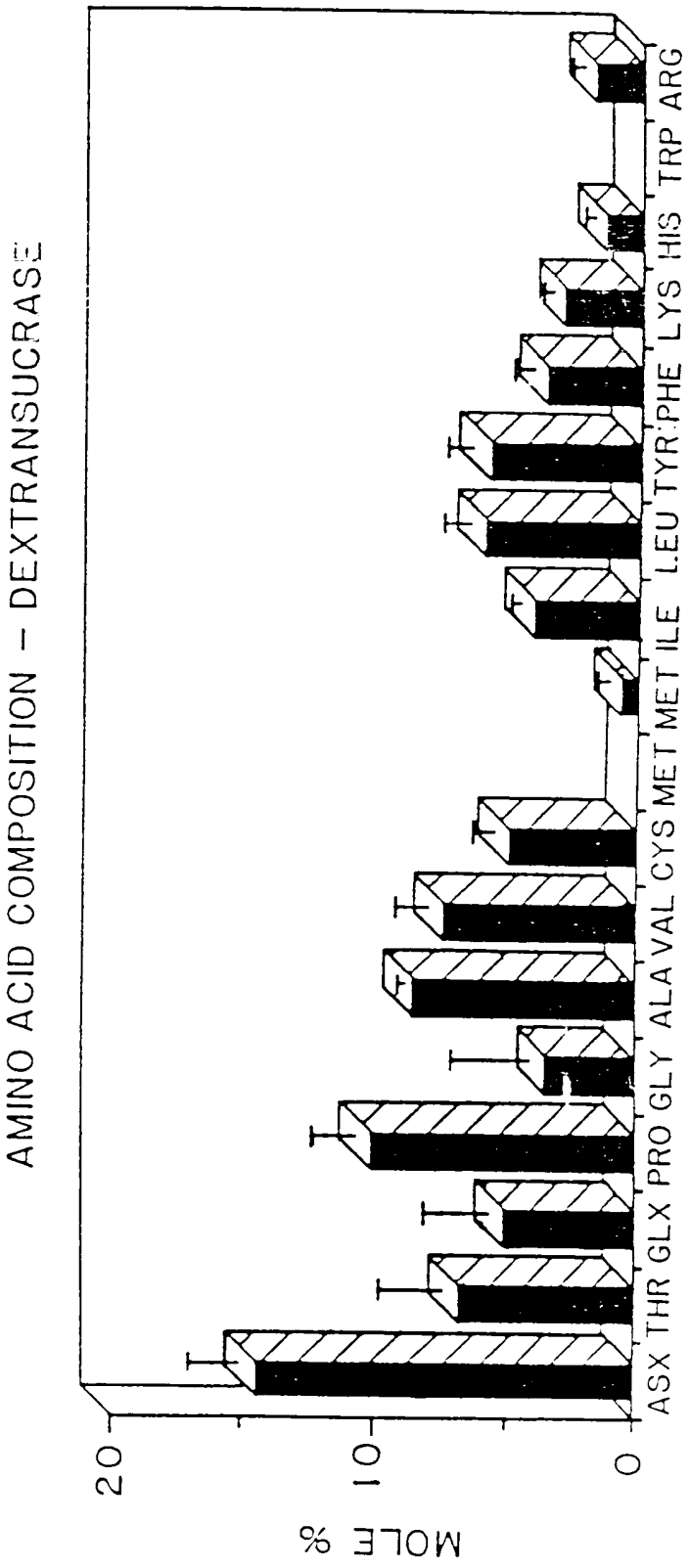

AMINO ACID COMPOSITION – DEXTRANSUCRASE
MOLE %
20
10
0
ASX THR GLX PRO GLY ALA VAL CYS MET ILE LEU TYR PHE LYS HIS TRP ARG
AMINO ACID

FIG. 5

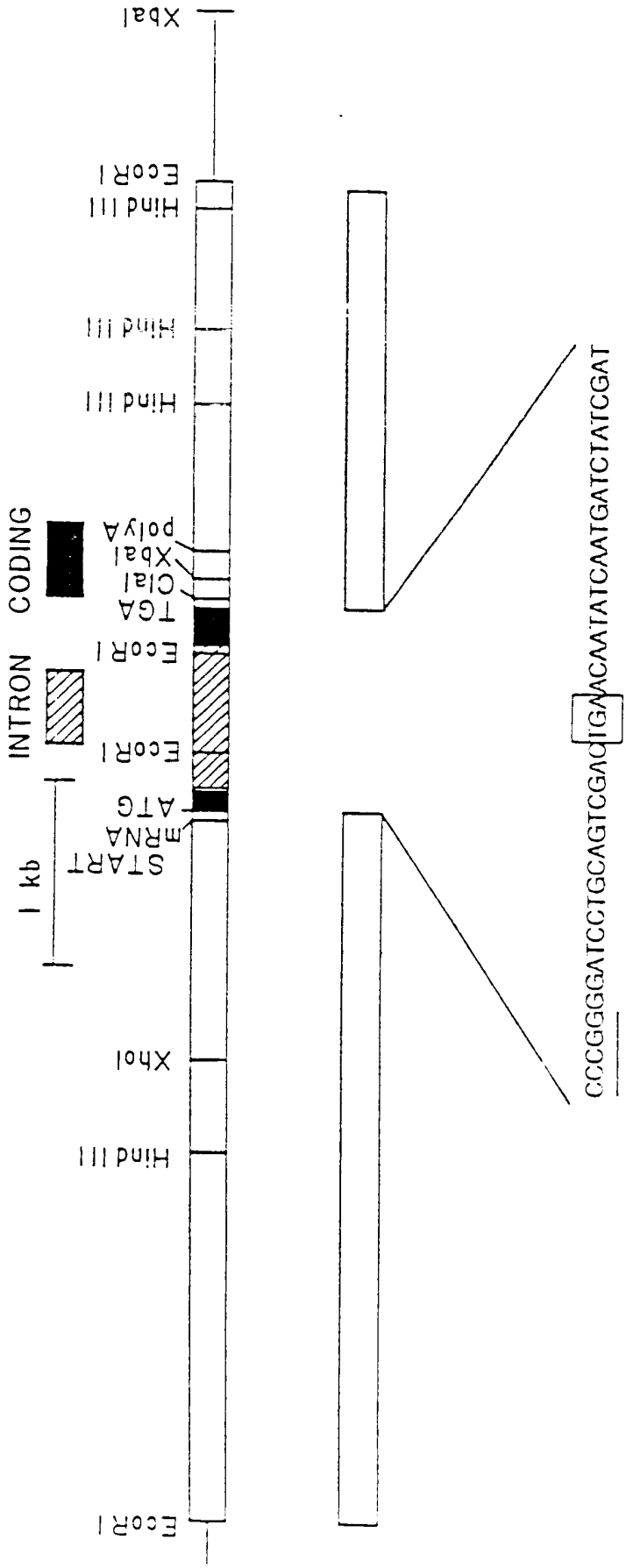

XbaI
EcoR I
Hind III
Hind III
Hind III
CODING
polyA
XbaI
ClaI
TGA
EcoR I
INTRON
EcoR I
ATG
mRNA
START
1 kb
XhoI
Hind III
EcoR I
CCCGGGGATCCTGCAGTCGACTGAGACAATATCAATGATCTATCGAT
Sma I
BamH I
Pst I
Sal I

FIG. 6

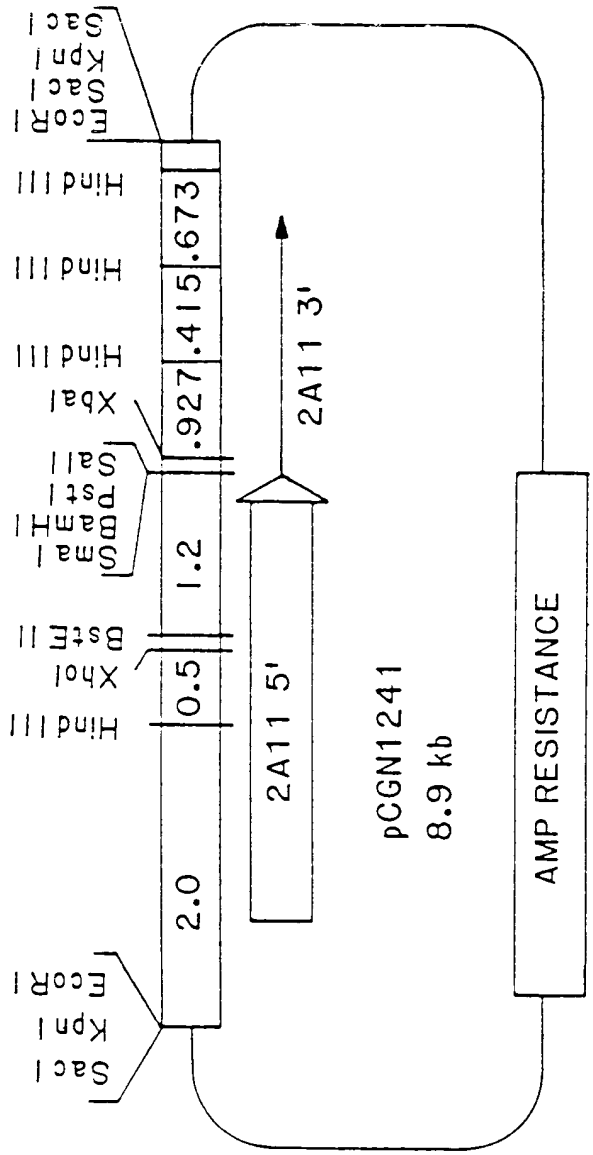

Sac I
Kpn I
Sac I
EcoR I
Hind III
Hind III
Hind III
Xba I
Sal I
Pst I
BamH I
Sma I
BstE II
Xho I
Hind III
2.0
0.5
1.2
.927
.415
.673
2A11 5'
2A11 3'
pCGN1241
8.9 kb
AMP RESISTANCE

FIG. 7

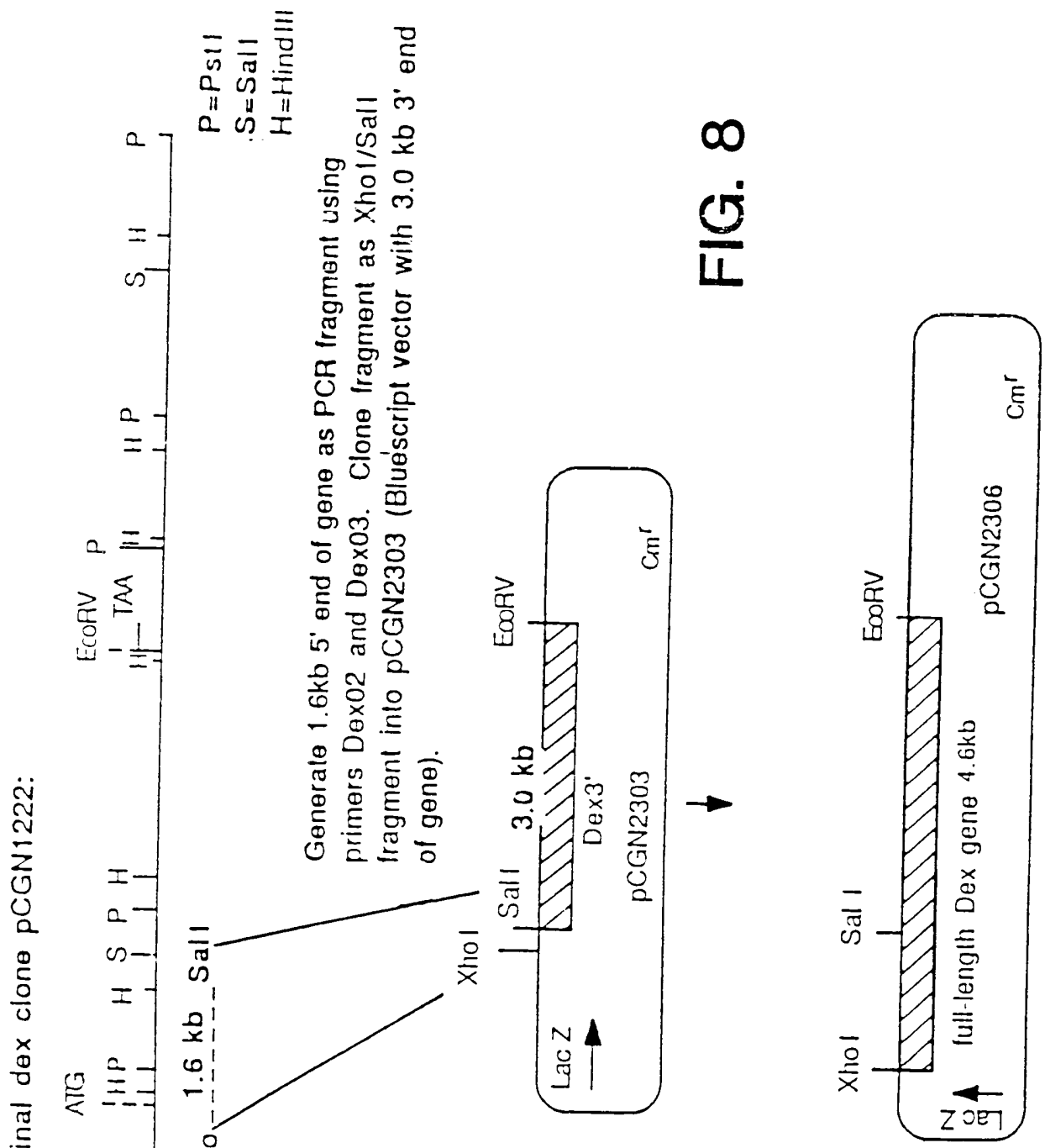
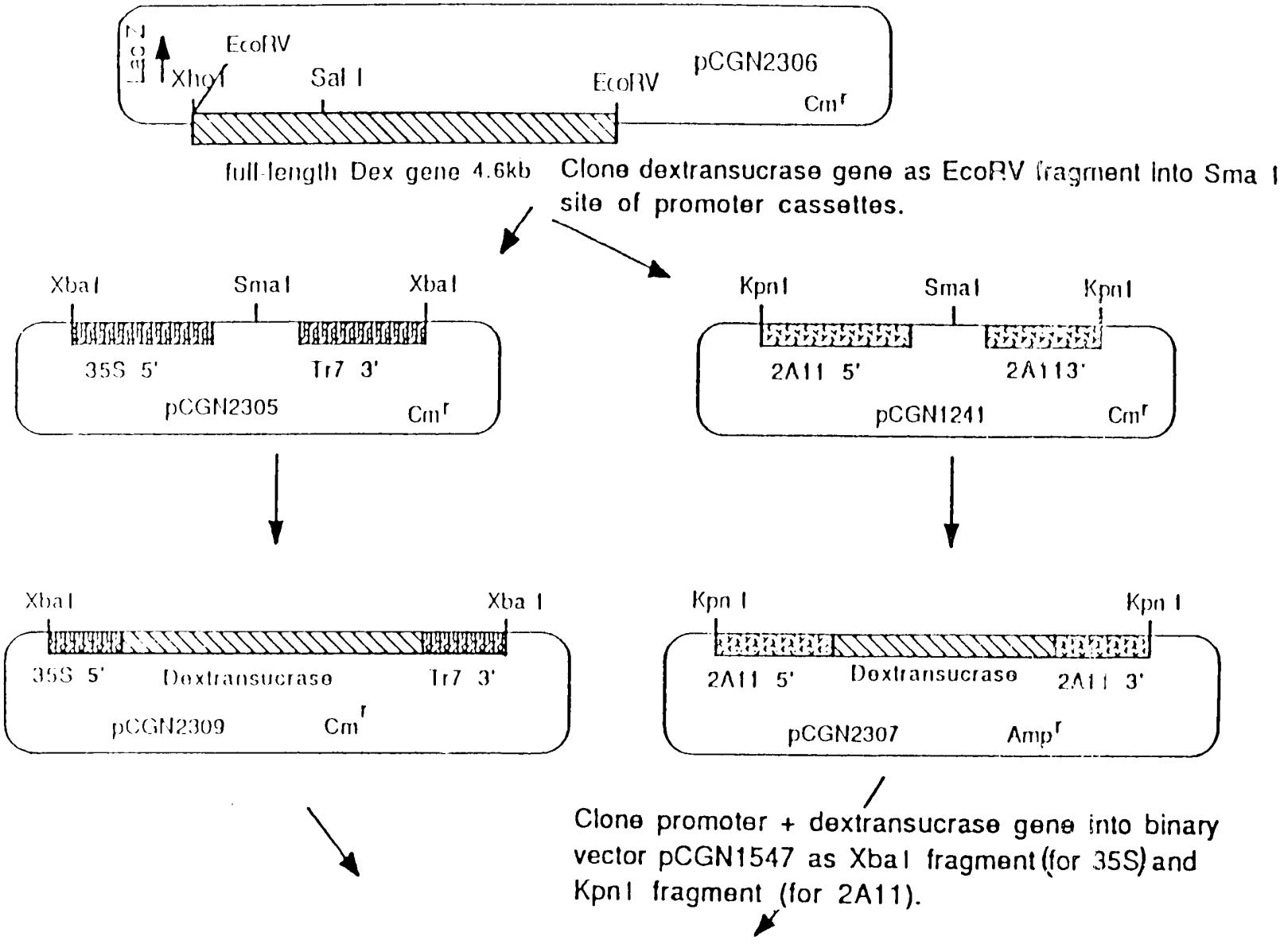

Insert from original dex clone pCGN1222:
ATG
EcoRV
TAA
P S S H S P H H P P S H P
P=PstI
S=SalI
H=HindIII
Xho
1.6 kb
SalI
Generate 1.6kb 5' end of gene as PCR fragment using primers Dex02 and Dex03. Clone fragment as XhoI/SalI fragment into pCGN2303 (Bluescript vector with 3.0 kb 3' end of gene).
XhoI
SalI
3.0 kb
Dex3'
EcoRV
pCGN2303
Cmr
Lac Z
XhoI
SalI
EcoRV
full-length Dex gene 4.6kb
pCGN2306
Cmr
Lac Z

FIG. 8

FIG. 9 -1


lacZ
EcoRV
Xho I
Sal I
EcoRV
pCGN2306
Cm^r
full-length Dex gene 4.6kb
Clone dextransucrase gene as EcoRV fragment into Sma I site of promoter cassettes.
Xba I
Sma I
Xba I
35S 5'
Tr7 3'
pCGN2305
Cm^r
Kpn I
Sma I
Kpn I
2A11 5'
2A11 3'
pCGN1241
Cm^r
Xba I
Xba I
35S 5'
Dextransucrase
Tr7 3'
pCGN2309
Cm^r
Kpn I
Kpn I
2A11 5'
Dextransucrase
2A11 3'
pCGN2307
Amp^r
Clone promoter + dextransucrase gene into binary vector pCGN1547 as Xba I fragment (for 35S) and Kpn I fragment (for 2A11).

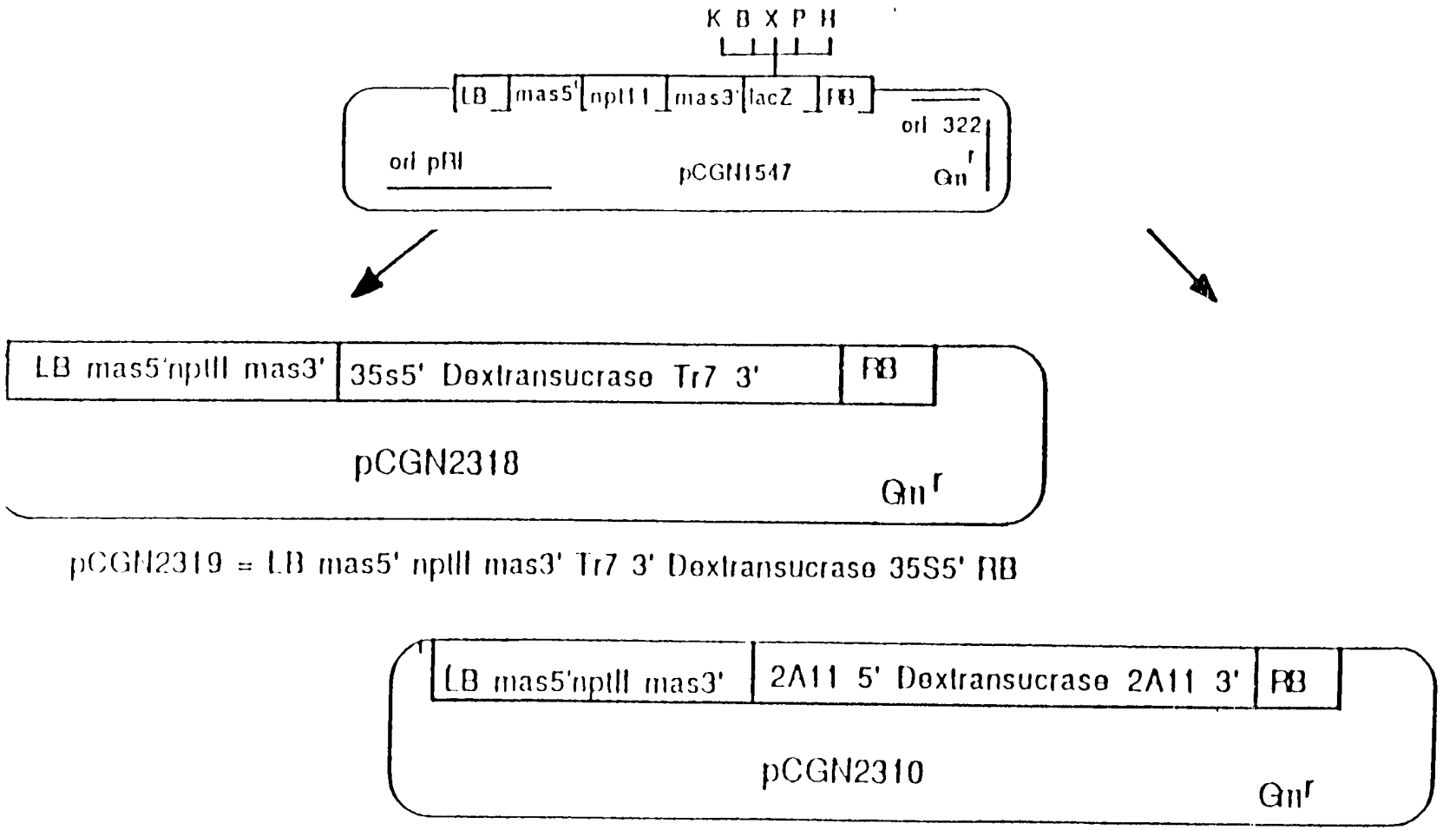

K B X P H
LB
mas5'
nptII
mas3'
lacZ
RB
ori 322
ori pRI
pCGN1547
Gm r
LB mas5'nptII mas3'
35s5' Dextransucrase Tr7 3'
RB
pCGN2318
Gm r
pCGN2319 = LB mas5' nptII mas3' Tr7 3' Dextransucrase 35S5' RB
LB mas5'nptII mas3'
2A11 5' Dextransucrase 2A11 3'
RB
pCGN2310
Gm r
pCGN2311 = LB mas 5' nptII mas3' 2A113' Dextransucrase 2A115' RB

FIG. 9-2

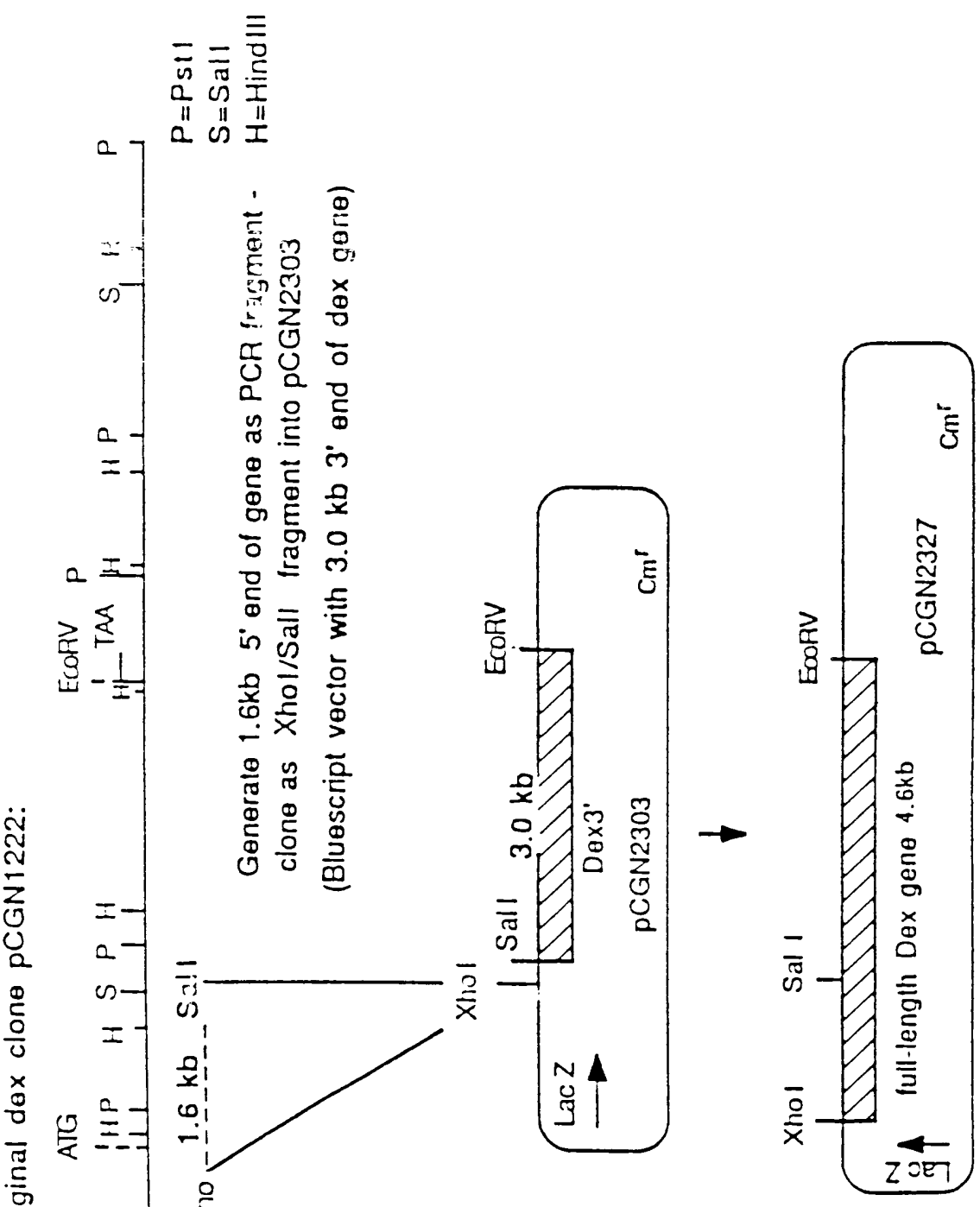
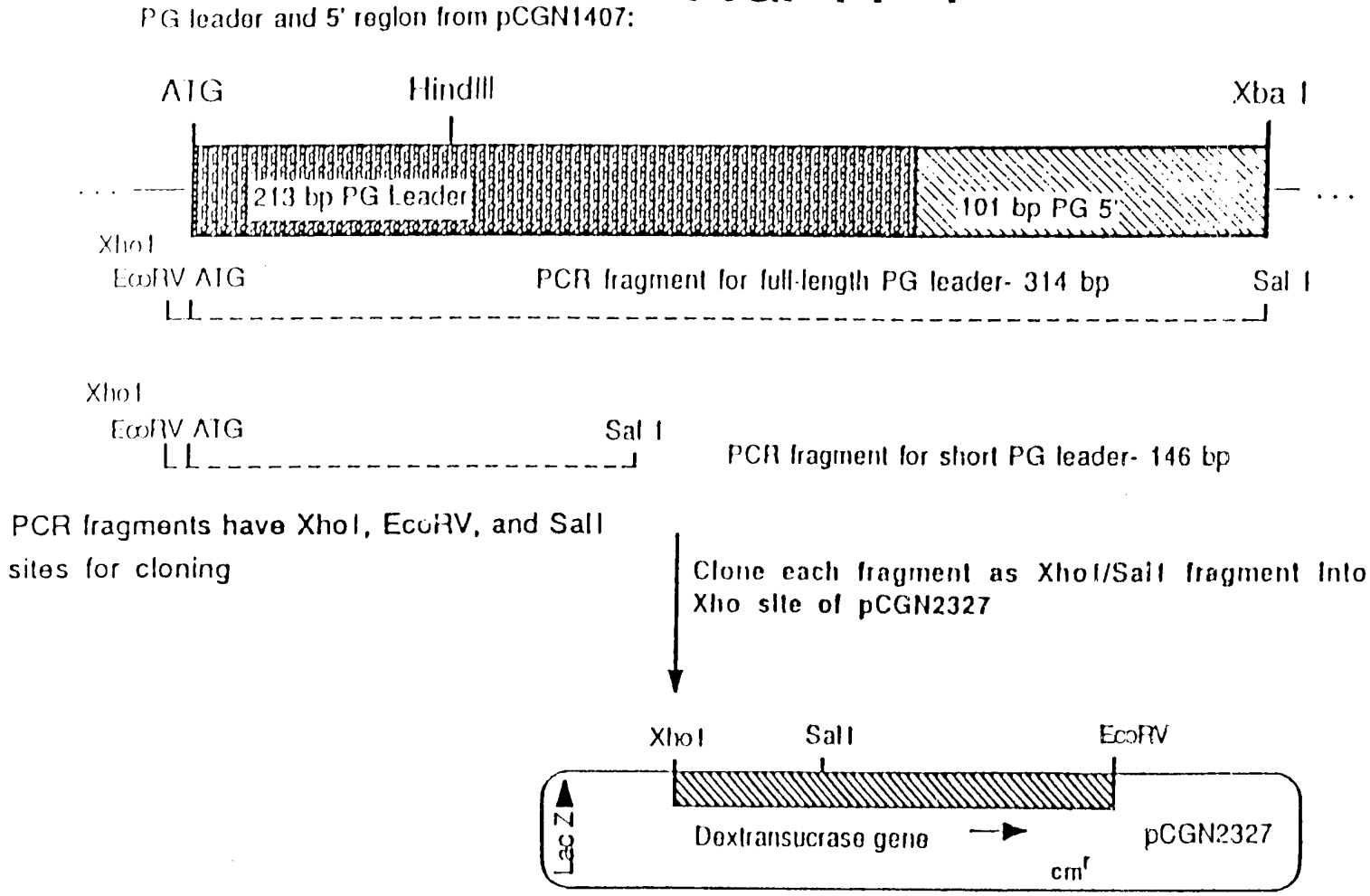

Insert from original dex clone pCGN1222:
ATG
EcoRV
TAA
P S S H P H S P H H P S H P
P=PstI
S=SalI
H=HindIII
Xho 1.6 kb SalI
Generate 1.6kb 5' end of gene as PCR fragment - clone as XhoI/SalI fragment into pCGN2303
(Bluescript vector with 3.0 kb 3' end of dex gene)
XhoI
SalI
3.0 kb
EcoRV
Dex3'
pCGN2303
Cmr
Lac Z
XhoI
SalI
EcoRV
full-length Dex gene 4.6kb
pCGN2327
Cmr
Lac Z

FIG. 10

FIG. 11-1


PG leader and 5' region from pCGN1407:
ATG
HindIII
Xba I
213 bp PG Leader
101 bp PG 5'
XhoI
EcoRV ATG
PCR fragment for full-length PG leader- 314 bp
Sal I
XhoI
EcoRV ATG
Sal I
PCR fragment for short PG leader- 146 bp
PCR fragments have XhoI, EcoRV, and SalI sites for cloning
Clone each fragment as XhoI/SalI fragment into Xho site of pCGN2327
XhoI
SalI
EcoRV
Lac Z
Dextransucrase gene
pCGN2327
cm^r

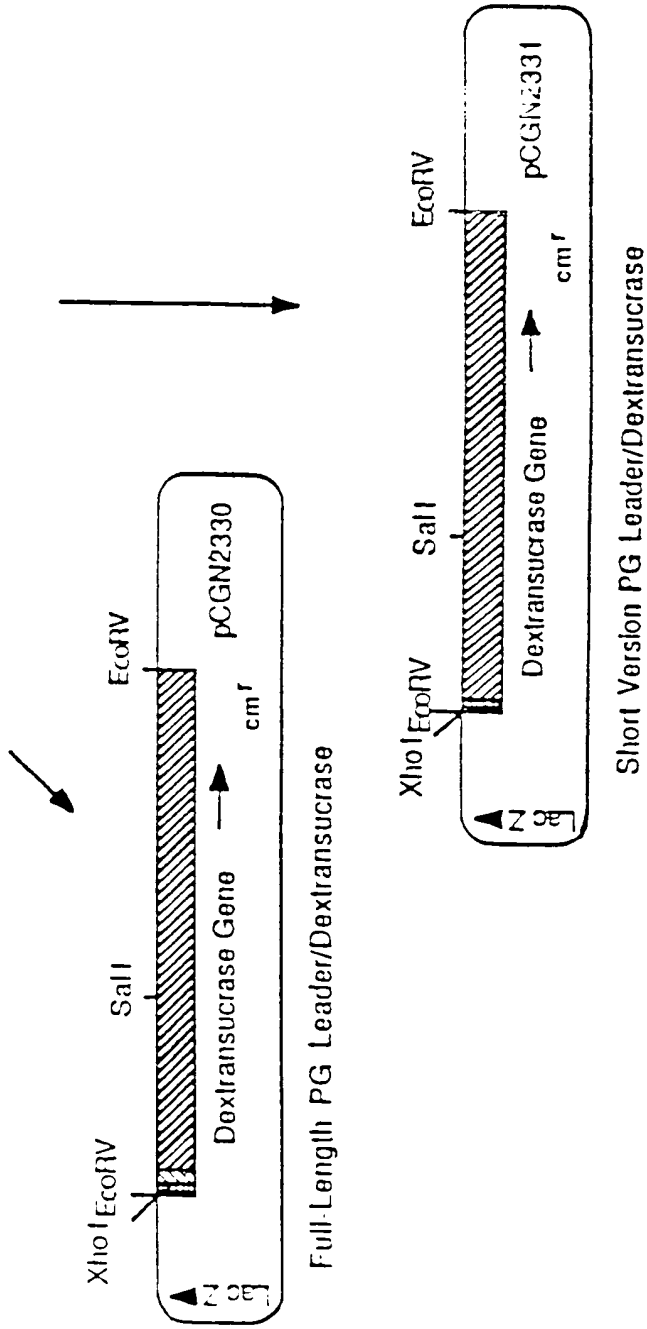
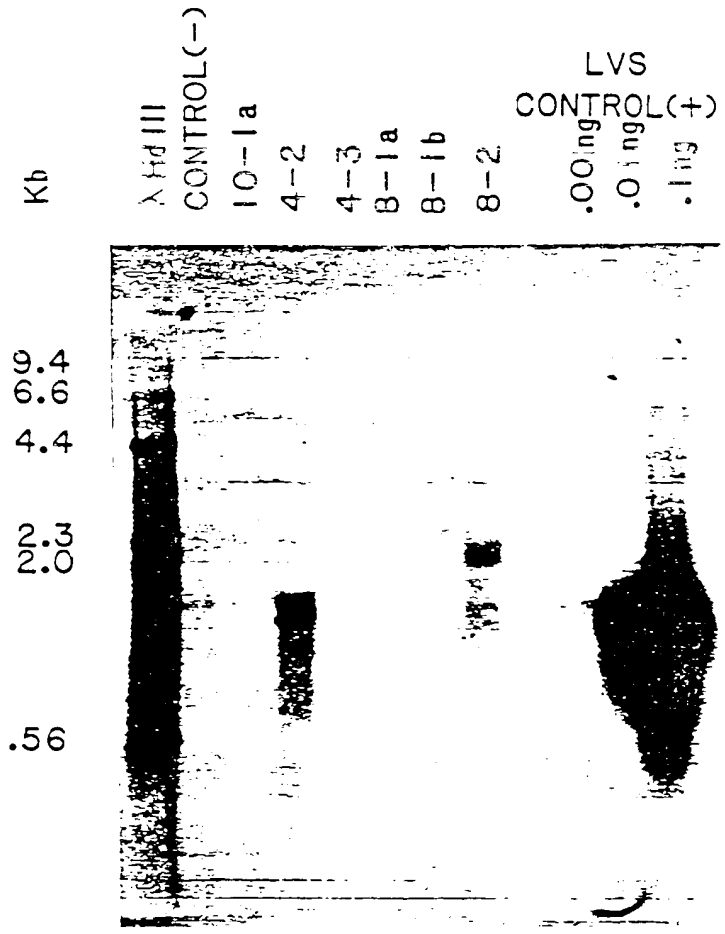

XhoI
EcoRV
SalI
EcoRV
Dextransucrase Gene
cm r
pCGN2330
Full-Length PG Leader/Dextransucrase
XhoI
EcoRV
SalI
EcoRV
Dextransucrase Gene
cm r
pCGN2331
Short Version PG Leader/Dextransucrase

FIG. 11-2

FIG. 12A



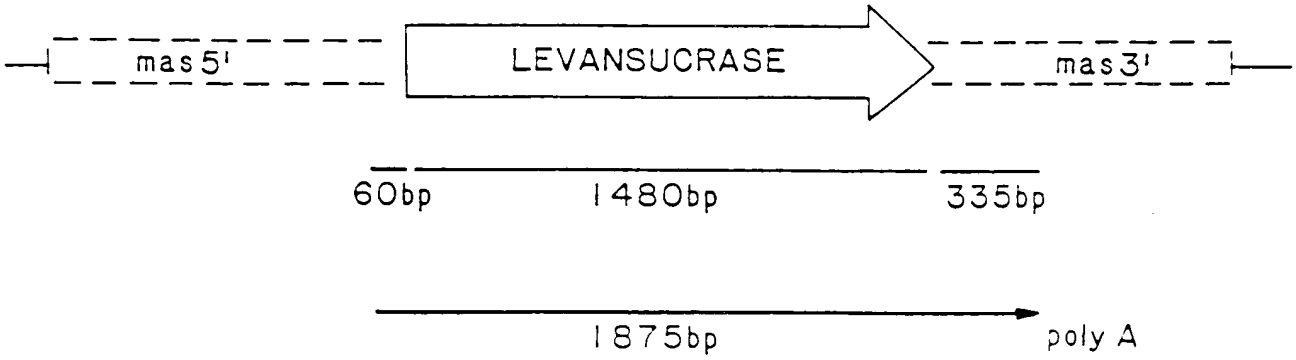


FIG. 12B

INTERNATIONAL SEARCH REPORT

International Application No PCT/US89/02729

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(4): A01H 1/04; C12N 1/00; C12N 5/00; C12N 15/00; C07H 21/04
U.S.CL.: 800/1; 435/172.1, 172.3, 240.4, 317.1, 320; 536/27

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
U.S.	536/27 435/172.1, 172.3, 240.4, 317.1, 320 800/1

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

USPTO Automated Patent System; DIALOG files: BIOTECH, PATENTS, WORLD PATENTS INDEX, See Attachment for search terms.

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
P,X,Y	US, A, 4,801,540 (HIATT) 31 January 1989, See entire document and page 4, lines 65-68.	1-30
A	Hortscience (St. Joseph, Michigan USA), Volume 13, Issued December 1978, Stevens et al, "Genetic potential for overcoming physiological limitations an adaptability, yield, and quality in the tomato", pages 673-678.	1-30
Y	Journal of Bacteriology (Washington DC, USA) Volume 153, Issued 1983, Gay et al., "Cloning structural Gene sacB, which codes for exenzyme levansucrase of Bacillus subtilis: Expression of the gene in E. coli", pages 1424-1431.	1-30

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search

September 1989

Date of Mailing of this International Search Report

30 OCT 1989

International Searching Authority

ISA/US

Signature of Authorized Officer

P. Rhodes

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

A	Theoretical and Applied Genetics (New York, USA) Volume 73, Issued 1987, Osborn et al, "Identification of restriction fragment length polymorphisms linked to genes controlling soluble solids content in tomato fruit", pages 350-356.	1-30

V ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers ____ because they relate to subject matter [2] not required to be searched by this Authority, namely:

2. ☐ Claim numbers ____, because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [3], specifically:

3. ☐ Claim numbers _____, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

SEE ATTACHMENT.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

Telephone Practice

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MR	Mauritania
BB	Barbados	GA	Gabon	MW	Malawi
BE	Belgium	GB	United Kingdom	NL	Netherlands
BF	Burkina Fasso	HU	Hungary	NO	Norway
BG	Bulgaria	IT	Italy	RO	Romania
BJ	Benin	JP	Japan	SD	Sudan
BR	Brazil	KP	Democratic People's Republic of Korea	SE	Sweden
CF	Central African Republic			SN	Senegal
CG	Congo	KR	Republic of Korea	SU	Soviet Union
CH	Switzerland	LI	Liechtenstein	TD	Chad
CM	Cameroon	LK	Sri Lanka	TG	Togo
DE	Germany, Federal Republic of	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		
ES	Spain	MG	Madagascar		